gedisa
editorial

Francisco J. Varela

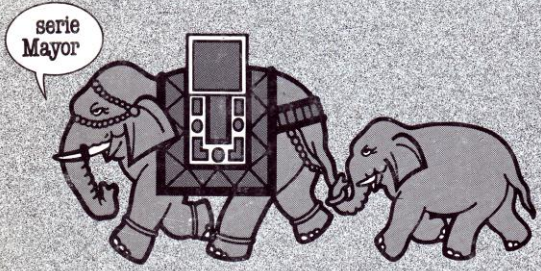
CONOCER

Las ciencias cognitivas: tendencias y perspectivas. Cartografía de las ideas actuales

La fase moderna de las ciencias cognitivas representa una mutación notable en la historia paralela de la mente y la naturaleza. Por primera vez la ciencia (es decir, el conjunto de científicos que definen qué debe ser la ciencia) reconoce plenamente la legitimidad de las investigaciones sobre el conocimiento mismo, en todos sus niveles, más allá de los límites tradicionalmente impuestos por la psicología o la epistemología.

Y por primera vez la sociedad occidental en su conjunto comienza a enfrentar en sus prácticas interrogantes tales como: ¿Es la mente una manipulación de símbolos? ¿Puede una máquina comprender el lenguaje? Se trata de preocupaciones que afectan la vida de la gente y no se limitan a ser teóricas. No es de extrañar que los medios manifiesten un constante interés por las ciencias cognitivas y la tecnología emparentada con ellas, ni que la inteligencia artificial haya penetrado profundamente en la mente de los jóvenes a través de los juegos de computación y la ciencia ficción. El efecto de esta fermentación es el

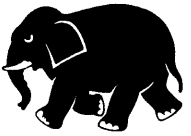
(sigue en la contratapa)



COLECCION EL MAMIFERO PARLANTE

Francisco J. Varela

CONOCER



COLECCION
EL MAMIFERO PARLANTE
Serie Mayor

CONOCER

Las ciencias cognitivas: tendencias y perspectivas. Cartografía de las ideas actuales

por

Francisco J. Varela

gedisa
editorial

Título del original en inglés:
Cognitive Science. A Cartography of Current Ideas.
© 1988 *by* Francisco J. Varela

Traducción: Carlos Gardini
Cubierta: Maqueta de colección: Julio Vivas
Realización: Gustavo Macri

Primera edición, Barcelona, 1990

Derechos para todas las ediciones en castellano

© by Editorial Gedisa
Muntaner, 460, entlo., 1ª
tel. 201 6000
08006 - Barcelona, España

ISBN: 84-7432-383-5
Depósito legal: B. 20.367 - 1990

Impreso en España
Printed in Spain

Impreso en Romanyà/Valls, S. A.
Verdaguer 1 - 08786 Capellades (Barcelona)

Queda prohibida la reproducción total o parcial por cualquier medio de impresión, en forma idéntica, extractada o modificada, en castellano o cualquier otro idioma.

ADVERTENCIA

ESTA ES UNA COPIA PRIVADA PARA FINES EXCLUSIVAMENTE EDUCACIONALES

QUEDA PROHIBIDA
LA VENTA, DISTRIBUCIÓN Y COMERCIALIZACIÓN

- El objeto de la biblioteca es facilitar y fomentar la educación otorgando préstamos gratuitos de libros a personas de los sectores más desposeídos de la sociedad que por motivos económicos, de situación geográfica o discapacidades físicas no tienen posibilidad para acceder a bibliotecas públicas, universitarias o gubernamentales. En consecuencia, una vez leído este libro se considera vencido el préstamo del mismo y deberá ser destruido. No hacerlo, usted, se hace responsable de los perjuicios que deriven de tal incumplimiento.
- Si usted puede financiar el libro, le recomendamos que lo compre en cualquier librería de su país.
- Este proyecto no obtiene ningún tipo de beneficio económico ni directa ni indirectamente.
- Si las leyes de su país no permiten este tipo de préstamo, absténgase de hacer uso de esta biblioteca virtual.

"Quién recibe una idea de mí, recibe instrucción sin disminuir la mía; igual que quién enciende su vela con la mía, recibe luz sin que yo quede a oscuras" ,

—Thomas Jefferson

COPIA PRIVADA

Para otras publicaciones visite
www.lecturasinegoismo.com
o puede escribirnos a
lecturasinegoismo@gmail.com
Referencia: 795

Editorial Gedisa ofrece
los siguientes títulos sobre

PSICOLOGIA COGNITIVA

pertenecientes a sus diferentes
colecciones y series
(Grupo "Psicología")

HILARY PUTNAM	*Representación y realidad*
DANIEL C. DENNETT	*La actitud intencional*
FRANCISCO J. VARELA	*Conocer*
JEROME BRUNER	*Realidad mental y mundos posibles*

Esta obra está dedicada,
con todo mi amor,
a mis hijos
Alejandra, Javier y Leonor

INDICE

INTRODUCCION 11

AGRADECIMIENTOS 19

1. **¿Por qué este ensayo?** 23
   Motivación 23
   Estructura de este ensayo 25

2. **Primera etapa. Los años jóvenes** 29
   Los comienzos 29
   La lógica y la ciencia de la mente 31
   Los frutos del movimiento cibernético 34

3. **Segunda etapa. Los símbolos: la hipótesis cognitivista** 37
   Los cognitivistas entran en escena 37
   Síntesis de la doctrina 38
   El vástago del cognitivismo: la inteligencia artificial 46
   Las ciencias cognitivas 48
   El procesamiento de la información en el cerebro 50
   Síntesis del disenso 52

4. **Tercera etapa. La emergencia: una alternativa ante la orientación simbólica** 53
   La autoorganización: las raíces de una alternativa 53
   La estrategia conexionista 57
   Emergencia y autoorganización 60

El conexionismo en la actualidad 67
Las emergencias neuronales 69
Los símbolos abandonan la escena 76

5. Cuarta etapa. La enacción: una alternativa ante la representación 87
Una insatisfacción más profunda 87
El redescubrimiento del sentido común 88
Las representaciones abandonan la escena 96
Ejemplos de enacción 101
Síntesis de la doctrina 108
Trabajar sin representaciones 110
Eslabones entre la emergencia y la enacción 113

6. Conclusiones 117

Introducción

Este pequeño libro se propone una tarea ambiciosa: presentar el panorama actual de las ciencias cognitivas, nombre con el que hoy designamos el análisis científico moderno del conocimiento en todas sus dimensiones. Este ensayo, pues, constituye una concisa visita guiada a un campo multidisciplinario que todavía no está bien definido.

Esta tarea no puede ser neutra: se tiene que emprender desde la perspectiva de alguien, preferentemente una parte interesada que hable de una actividad de la cual participa. Como pronto descubrirá el lector, mi perspectiva, desarrollada al cabo de unos veinte años de investigación, está expuesta explícitamente en el texto. No la repetiré aquí, pero caben dos observaciones.

Ante todo, una observación epistemológica. Cada época de la historia humana produce, a través de sus prácticas sociales cotidianas y

su lenguaje, una estructura imaginaria. La ciencia forma parte de estas prácticas sociales, y las ideas científicas acerca de la naturaleza constituyen apenas una dimensión de esta estructura imaginaria. Los historiadores y filósofos modernos, desde Alexander Koyré, han demostrado que la imaginación científica sufre mutaciones radicales de una época a otra, y que la ciencia se parece más a una epopeya novelística que a un progreso lineal. La historia humana de la naturaleza es una narración que merece ser contada de más de un modo.

Lo que resulta menos evidente es que dicha historia humana de la naturaleza se corresponde con una historia de las ideas sobre el conocimiento de sí mismo. Así, la *physis* griega y el método socrático, o los ensayos de Montaigne y la temprana ciencia francesa, son pares interdependientes. Reflejándose mutuamente, el sí-mismo y la naturaleza se desplazan en el tiempo como una pareja de bailarines. En Occidente aún no se ha escrito la historia natural del conocimiento del sí-mismo[1], pero es justo aclarar que siempre hubo precursores de lo que hoy llamamos ciencias cognitivas, en la medida en que la mente humana es la fuente primordial y el ejemplo más accesible de la cognición y el conocimiento.

La fase moderna de las ciencias cognitivas

[1] Véase, sin embargo, F. Schuarzman, *Teoría de la expresión*, Edit. Universitaria, Santiago, 1968.

representa una mutación notable en esta historia paralela de la mente y la naturaleza. Por primera vez la ciencia (es decir, el conjunto de científicos que definen qué debe ser la ciencia) reconoce plenamente la legitimidad de las investigaciones sobre el conocimiento mismo, en todos sus niveles, más allá de los límites tradicionalmente impuestos por la psicología o la epistemología. Esta mutación, que tiene apenas treinta años, fue enfáticamente introducida mediante el programa cognitivista comentado en el texto, tal como el programa darwiniano inauguró el estudio científico de la evolución, aunque otros se habían interesado antes en ella.

Más aun, a través de esta mutación, el conocimiento se ha ligado tangiblemente a una tecnología que *transforma* las prácticas sociales que lo posibilitaron: la inteligencia artificial constituye el ejemplo más visible. La tecnología, entre otras cosas, actúa como un amplificador. No podemos separar las ciencias cognitivas de la tecnología cognitiva sin despojar a una u otra de un vital elemento complementario. En otras palabras, a través de la tecnología, la exploración científica de la mente brinda a la sociedad un inadvertido espejo de sí misma que trasciende el círculo del filósofo, el psicólogo o el pensador.

Y por primera vez la sociedad occidental en su conjunto comienza a enfrentar en sus prácticas interrogantes, tales como: ¿Es la mente una manipulación de símbolos? ¿Puede una máquina comprender el lenguaje? Se trata de preocupa-

ciones que afectan la vida de la gente y no se limitan a ser teóricas. No es de extrañar que los medios manifiesten un constante interés por las ciencias cognitivas y la tecnología emparentada con ellas, ni que la inteligencia artificial haya penetrado profundamente en la mente de los jóvenes a través de los juegos de computación y la ciencia ficción. El efecto de esta fermentación es el siguiente: mientras durante milenios las gentes tuvieron una comprensión espontánea de sí mismas, según la cultura de su época, por primera vez esta visión popular de la mente entra en contacto con la ciencia y es transformada por ella. Muchos deplorarán esta revolución, mientras que otros la celebrarán.

Sea como fuere, este fenómeno adquiere creciente celeridad. El fecundo diálogo entre investigadores, tecnólogos y público encierra un potencial para la transformación de la conciencia humana que yo encuentro fascinante, pues se trata de una de las más interesantes aventuras que hoy enfrentamos. Este texto constituye un modesto pero —espero— significativo aporte a este diálogo transformador. Por ejemplo, en este ensayo cuestionaré la difundida idea de que el conocimiento está relacionado con el procesamiento de información. Argumentaré que la información es semejante a un moderno flogisto que intenta explicar la estructura del conocimiento apoyándose en un orden de cosas preexistente. El pivote de la cognición es precisamente su capacidad para explicar la significación y las regu-

laridades; la información no debe aparecer como un orden intrínseco sino como un orden emergente de las actividades cognitivas mismas. Si ello se verifica, nuestra ingenua comprensión de las relaciones que entablamos con el mundo cambiará drásticamente.

El segundo punto es de índole sociopolítica. Como a cualquiera que haya examinado de cerca una disciplina científica, las ciencias cognitivas me han parecido un mosaico de perspectivas más o menos compatibles antes que un dominio homogéneo. No obstante, es obvio que la ciencia, en cuanto actividad social, está atravesada por corrientes de poder que infunden más autoridad a ciertas voces que a otras. Europa fue epicentro de la ciencia hasta la época de las guerras mundiales, pero es indiscutible que dicho papel corresponde hoy a los Estados Unidos. De hecho, la mayoría de mis colegas europeos y sudamericanos consideran que los Estados Unidos constituyen la vara para medir la calidad, el prestigio y el mérito.

Esto es aun más cierto en el dominio de las ciencias cognitivas, en la medida en que la revolución cognitiva moderna recibió gran influencia de las investigaciones realizadas en la Costa Este de los Estados Unidos, sobre todo en el MIT. Esta presunta tradición cognitivista se ha transformado en la ortodoxia de la comunidad científica, como veremos en la primera parte de este en-

sayo. Al leer a los portavoces de esta tradición se tiene la impresión de que (a) el enfoque computacional es el único modo de abordar la ciencia cognitiva, y (b) nada que sea anterior al surgimiento de esta tecnología merece conservarse, ni siquiera de manera transformada.

Desde luego, no es sólo un enunciado científico sino político, pues brinda pautas para juzgar el trabajo científico: relevante es aquello que yo considero relevante. Confieso ser un inveterado amante de la heterodoxia, y un ávido buscador de la diversidad. Desde luego, es preciso conocer bien las importantes tesis y resultados producidos por las predominantes ciencias cognitivas estilo MIT, pero no puedo aceptar que los enfoques tradicionales sean los únicos válidos.

Esta cuestión resulta especialmente delicada en Europa, pues durante el siglo veinte hubo muy importantes aportes al trasfondo conceptual de lo que hoy llamamos ciencias cognitivas, y con frecuencia se los ignora por completo. Pienso ante todo en: a) el movimiento fenomenológico, especialmente Edmund Husserl y Maurice Merleau-Ponty, y b) los enfoques inaugurados por Jean Piaget en epistemología genética. Estas escuelas exploraron profundos interrogantes cognitivos y sus mecanismos, como luego comentaremos en este texto. Lamentablemente, han estado casi totalmente ausentes de la ortodoxia cognitivista, y en consecuencia sus ideas básicas a veces se reinventan como si fueran novedades. Un notable ejemplo es el redescubri-

miento de la intencionalidad en los estudios cognitivos, algo claramente expresado por pensadores europeos en la década de 1940, pero totalmente ignorado en la ciencia cognitiva tradicional hasta 1980.

Si Europa ha de participar en pie de igualdad en el diálogo destinado a desarrollar la nueva ciencia de la mente, es esencial que se valga de sus singulares tradiciones, que se anticipe a los planteos y que permanezca abierta a diversos estilos de trabajo. Es interesante señalar que el Japón ya se desplaza enérgicamente en esa dirección; en este campo, más que en ningún otro, se siente la presencia del Japón en la ciencia internacional, un hecho directamente reflejado en su liderazgo tecnológico. El lector verá que guardo una desembozada distancia frente a la ciencia cognitiva tradicional, e insisto en las diferencias existentes. Lo hago, por cierto, a partir de convincentes razones científicas, pero también se trata de una posición tomada en lo relacionado con la sociología de la ciencia.

Agradecimientos

Escribí estas páginas a requerimiento de la Royal Dutch Shell Corporation. A fines de 1985 el Departamento de Planificación me invitó a dar una conferencia acerca de mi labor en el Shell Centre de Londres. Cuando manifesté mi sorpresa ante esta invitación, recibí una interesante respuesta: Shell necesita comprenderse a sí misma como sistema complejo de aprendizaje. Inmediatamente sentí la tentación de aceptar. Conocí a un grupo inteligente que no sólo me ayudó a aclarar mis propias ideas, sino que me permitió comprender que lo que parece abstruso y distante en la investigación puede desencadenar cambios de mentalidad con consecuencias muy concretas. Como resultado de nuestro diálogo, me solicitaron que redactara un panorama actualizado de las ciencias y tecnologías de la cognición desde mi punto de vista, enfatizando las que yo consideraba tendencias emergentes. El espinazo de este libro es el texto de ese informe, titulado *Ciencia y tecnología de la cognición:*

tendencias emergentes, salvo por pasajes que he omitido porque respondían a necesidades específicas de la Shell, y extensos añadidos y actualizaciones destinadas a la publicación en Francia. Deseo agradecer a la Shell su respaldo e interés, sobre todo a Peter Schwartz, ex director del Business Environment Group del Shell Centre, y a Arie de Geus, director de Planificación, sin quienes este trabajo jamás habría llegado a buen término.

También deseo expresar mi gratitud a mis colegas de París por su calurosa recepción en Francia, que me permitió continuar mi labor cuando me resultaba imposible en mi patria, Chile, asolada por la peste del fascismo. Mi particular agadecimiento a CREA (Centre Recherche Epistemologie Appliqué) de la École Polytechnique, donde aún prospera el aspecto interdisciplinario de la investigación cognitiva. Mi especial gratitud a Jean-Pierre Dupuy, director y fundador de CREA, por muchos años de amistad y entusiasmo hacia mi labor. También agradezco al Institut des Neurosciences (CNRS-Université Paris VI) por brindarme espacio de laboratorio y recibirme como miembro. Mi especial agradecimiento a Michel Imbert por su apertura mental y su respaldo.

Mi gratitud, en fin, a la Fondation de France por darme una cátedra de ciencias de 1986 a 1990, y al Prince Charitable Fund por sus subvenciones científicas. Sin tal respaldo, estas páginas no existirían.

Fragmentos de este texto se han publicado con estos títulos:

Varela, F., "The science and technology of cognition: Emergent directions", en J. L. Ross (comp.), *Economics and Artificial Intelligence*, Primer Simposio Nacional del IFAC, Nueva York, Pergamon Press, 1987, págs. 1-9.

Varela, F. (1988), "Perception and the origin of cognition: A Cartography of current ideas", en F. Varela y J. P. Dupuy (comps.), *Understanding Origin: Ideas on Origins of Life, Mind and Society*, Stanford International Symposium (de inminente publicación).

Parte del material de los capítulos 3, 4 y 5 ha sido adaptado de un trabajo actualmente en curso:

F. Varela y E. Thompson, *Worlds without Grounds: Cognitive Science and Human Experience* (inédito).*

* La presente traducción se realizó a partir del original inglés y se cotejó con la versión francesa, *Connaître: les sciences cognitives*, Editions du Seuil, 1988 (trad. P. Lavoie), París que constituye la primera publicación de este trabajo en forma de libro. *(N. del T.)*

1

¿Por qué este ensayo?

Motivación

Las ciencias y tecnologías de la cognición (CTC) constituyen la revolución conceptual y tecnológica más significativa desde la física atómica, pues ejercen un impacto de largo plazo en todos los niveles de la sociedad. La tecnología de la información (TI) es sólo el aspecto más visible de este vasto complejo de investigaciones y aplicaciones cuyos principales intereses se orientan hacia el conocimiento, la información y la comunicación.

Las CTC son un híbrido de diversas disciplinas interrelacionadas, y cada cual aporta sus intereses y preocupaciones propios (Figura 1). Se puede argumentar que la ciencia cognitiva es una disciplina aparte cuyo polo tecnológico es la inteligencia artificial (AI), y que no se las debería confundir como hacemos aquí. No obstante, el distingo entre ciencia aplicada y ciencia pura parece fuera de lugar en esta empresa (como para

la mayoría de las ciencias modernas: pensemos en la biotecnología). Uno de nuestros propósitos es demostrar que concentrarse *sólo* en el aspecto tecnológico o el aspecto "puro" de las CTC equivale a perder de vista su vitalidad y su futuro. Este campo resulta fascinante precisamente porque conjuga puntos de vista que proceden de fuentes alejadas entre sí como son la ingeniería informática y el pensamiento filosófico. Es fácil comprenderlo cuando se enumeran algunas de las áreas que interesan al científico cognitivo de hoy: la percepción, el lenguaje, la inferencia y la acción. Dichos intereses también se reflejan en las siguientes tecnologías: reconocimiento de

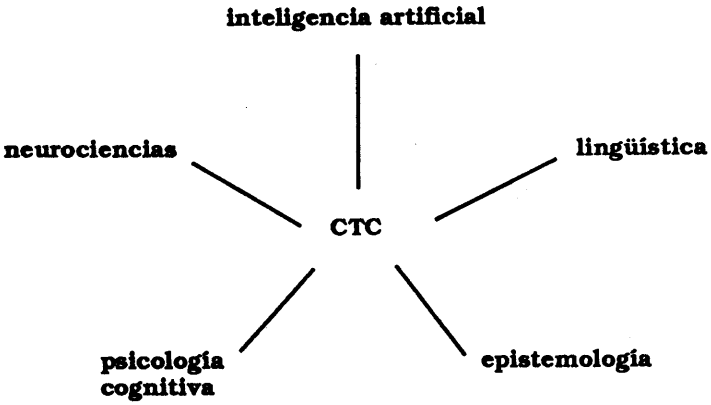


Figura 1. Las principales disciplinas que forman parte de las CTC.

imágenes, comprensión del lenguaje, síntesis de programas, robótica.

Las CTC tienen poco más de 40 años. No están afianzadas como ciencias maduras que disponen de un rumbo preciso y una numerosa comunidad de investigadores, como ocurre, por ejemplo, con la física atómica o la biología molecular. Por lo tanto, el desarrollo futuro de las CTC dista de ser claro, pero sus productos ya han ejercido un profundo impacto, y muchos opinan que seguirán ejerciéndolo. Pero el progreso en este campo se basa en audaces apuestas conceptuales: algo parecido a tratar de enviar un hombre a la Luna sin saber dónde queda la Luna.

Estructura de este ensayo

El propósito de este ensayo consiste en brindar una radiografía de la situación *actual* de las CTC. Lo haremos dividiendo las CTC en cuatro etapas o capas conceptualmente distintas que han surgido en forma más o menos sucesiva en los últimos 40 años. Las cuatro etapas son las siguientes:

Primera etapa. Los años fundacionales (1943-53).

Segunda etapa. Los símbolos: el paradigma cognitivista.

Tercera etapa. La emergencia: una alternativa ante la manipulación de símbolos.

Cuarta etapa. La enacción:* una alternativa ante la representación.

A través de este abordaje en cuatro etapas examinaremos la base de lo que ya está establecido como un "paradigma"[3] claramente perfilado (etapas primera y segunda), y el hecho esencial de que este paradigma establecido está allanando el camino a nuevas perspectivas emergentes (etapas tercera y cuarta). Estas provocativas heterodoxias encierran el potencial para causar un profundo impacto en el futuro.

Esta evaluación de las tendencias emergentes se propone indicar el peligro de un excesivo predominio de la orientación prevaleciente en las CTC, sin que se otorgue al menos cierto espacio a otros enfoques. En la medida en que cada grupo social ha tomado, y continuará tomando, importantes decisiones acerca de cuáles CTC son más apropiadas y compatibles con su identidad, las citadas posibilidades merecen una seria reflexión. En el futuro inmediato, estas posibilidades podrían inducir significativos cambios en áreas, tales como la automatización de las oficinas y el análisis de imágenes. Dentro de cinco o diez

* La elección de este término se justifica en el capítulo 5. [N. del T. francés.]

[3] Un paradigma, según T. S. Kuhn (*The Structure of Scientific Revolutions*, 2ª ed., Chicago University Press, 1970) es el modo habitual de referirse al conjunto coherente de ideas científicas que se aceptan como explicación de un corpus fenomenológico. La idea es discutible, pero eso no nos interesa aquí.

años, podrían conducir a un modo totalmente diferente de incorporar las CTC a nuestras preocupaciones sociales por el funcionamiento y la identidad. Por ejemplo, ¿debemos dar prioridad a la compatibilidad entre ordenadores o a la construcción de máquinas capaces de comprender el lenguaje humano? Como veremos, quien enfrenta una decisión de este tipo —trátese de un rector universitario o del Ministerio de Investigación y Tecnología— debe tener en cuenta todas las posibilidades que presentan las CTC: un mero análisis de costes económicos no será suficiente, pues se pagará un precio muy alto si por ignorancia se ahonda la inercia de ciertas estructuras transformándolas en escollo para cambios futuros.

2

Primera etapa. Los años jóvenes

Los comienzos

Comenzaremos con los años formativos de las CTC, que abarcan el período 1940-1956. Esta breve ojeada a las raíces es necesaria. Una ciencia que olvida su pasado está condenada a repetir sus errores, y es incapaz de evaluar su desarrollo. Por cierto, nuestra breve excursión no se propone ser una historia exhaustiva, sino sólo tocar aquellos problemas de relevancia directa para el presente ensayo.[4]

[4] Este capítulo debe mucho a nuestro reciente trabajo colectivo sobre la historia temprana de la cibernética y la autoorganización, publicada como *Cahiers du CREA*, Nos. 7-9, y en particular al artículo de J. P. Dupuy, "L'essor de la première cybernetique", Nº 7, págs. 7-140. Otra fuente útil es S. Heims, *John Von Neumann and Norbert Wiener* (MIT Press, 1980). El reciente libro de H. Gardner, *The Mind's New Science: A History of the Cognitive Revolution* (Nueva York: Basic Books, 1985) comenta este período sólo de manera superficial.

De hecho, casi todos los temas que hoy se debaten activamente se introdujeron en esos años de formación, lo cual evidencia que los problemas son profundos y difíciles de abordar. Los "padres fundadores" sabían muy bien que sus preocupaciones conducían a una nueva ciencia, y la bautizaron con un nombre que aludía explícitamente a su orientación *epistemológica*.[5] Esta palabra ha caído en desuso, y en la actualidad muchos científicos cognitivos ni siquiera reconocerían el parentesco. No mencionamos esto porque sí. Refleja el hecho de que para consolidarse como ciencia, en su clara orientación cognitivista (la segunda etapa, en este texto), la futura ciencia cognitiva tuvo que cercenar sus complejas raíces, más difusas pero también más ricas. Ello ocurre a menudo en la historia de la ciencia: es el precio que se paga por pasar de una etapa exploratoria a la constitución de un paradigma: de la nube al cristal.

Los años pioneros fueron el resultado de un intenso diálogo entre gentes de muy diversa formación: un esfuerzo interdisciplinario singularmente feliz que se produjo con notable coincidencia en Europa y los Estados Unidos. En Suiza Jean Piaget formuló un programa de investigación

[5] Usamos la palabra en su connotación anglosajona, relacionada con la teoría del conocimiento, no en la acepción a menudo utilizada en historia/filosofía de las ciencias en Francia, tal como la popularizaron estudiosos como Gaston Bachelard.

en lo que él denominaba epistemología *genética*, mientras Konrad Lorenz describía su visión de una epistemología *evolutiva*. Al mismo tiempo, en los Estados Unidos, Warren McCulloch empezaba a hablar de epistemología *experimental*.

Esta sincronicidad de los esfuerzos para naturalizar la epistemología tuvo su desarrollo más intenso en una zona geográfica localizada, principalmente centrada alrededor del MIT y de Princeton. Los principales actores fueron un puñado de científicos respaldados por algunas mentes pródigas, principalmente John von Neumann, Norbert Wiener, Alan Turing y Warren McCulloch. Es interesante recordar que estos esfuerzos cobraron forma visible al amparo del neologismo acuñado por Wiener: *cibernética*.[6]

La lógica y la ciencia de la mente

La intención expresa del movimiento cibernético se puede resumir en pocas palabras: crear una *ciencia de la mente*. Sus líderes opinaban —aunque desde luego las diferencias filosóficas variaban sustancialmente— que el estudio de los fenómenos mentales había estado demasiado

[6] La mejor fuente para esto son las muy citadas Conferencias de Macy, publicadas como *Cybernetics — Circular causal and feedback mechanisms in biological and social systems* (Nueva York: Josiah Macy Jr. Foundation, 5 volúmenes).

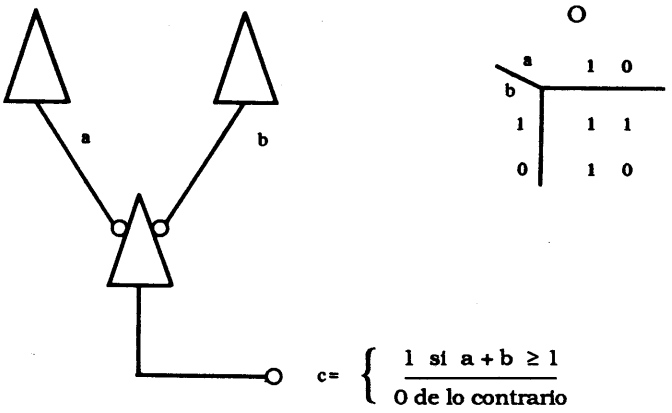


Figura 2. Tres neuronas de McCulloch-Pitts conectadas para realizar la opoeración lógica O. La neurona c se activa cuando una (o ambas) de las otras neuronas están activas, una conducta idéntica a la tabla de verdad de la operación O.

tiempo en manos de psicólogos y filósofos, y anhelaban expresar los procesos que subyacían a los fenómenos mentales en mecanismos explícitos y formalismos matemáticos.

Uno de los mejores ejemplos de este modo de pensar fue el trabajo seminal de McCulloch y Pitts de 1943, titulado: "Un cálculo lógico inmanente en la actividad nerviosa".[7] Este artículo dio varios pasos importantes. Primero, proponer que la *lógica* es la disciplina adecuada para com-

[7] "A logical calculus immanent in nervous activity", *Bulletin of Mathematical Biophysics*, vol. 5, 1943.

prender el cerebro y la actividad mental. Segundo, ver que el cerebro es un dispositivo que *encarna* principios lógicos en sus elementos constitutivos o neuronas. Se considera que cada neurona es un autómata-umbral cuyo estado, activo o inactivo, indicaría un valor lógico de verdad o falsedad, según el caso. Tales neuronas se podrían conectar entre sí, y sus interconexiones desempeñarían el papel de las operaciones lógicas ("y", "no", y demás). A partir de allí, se podía considerar al cerebro entero como una máquina deductiva.

En manos de John von Neumann de Princeton, estas ideas constituirían un paso fundamental para el invento del ordenador digital.[8] En esa época usó tubos de vacío para representar las neuronas McCulloch-Pitts. Hoy encontramos *chips* de silicio en su lugar, pero los ordenadores modernos (con la excepción de las muy modernas máquinas paralelistas) todavía respetan la *arquitectura von Neumann*. Este nombre tan técnico se ha convertido en una idea casi cotidiana con la llegada de los microordenadores y su familiar configuración: una unidad procesadora central, una memoria, y una unidad para operaciones aritméticas, todas interconectadas por un *bus* portador de señales. Se trató por

[8] Para una interesante perspectiva acerca de este momento histórico/conceptual, véase también A. Hodges, *Alan Turing: The Enigma of Intelligence* (Nueva York: Touchstone, 1984).

cierto de un decisivo avance tecnológico, pero su mayor importancia consistió en echar los cimientos para el enfoque dominante en el estudio científico de la mente, el cual cristalizaría en la siguiente década como el paradigma cognitivista (Etapa 2).

Los frutos del movimiento cibernético

Más allá del trabajo de McCulloch-Pitts y sus tangibles consecuencias, la fase cibernética de las CTC produjo una asombrosa gama de resultados, al margen de su influencia de largo plazo (a menudo subterránea). He aquí algunos:

— la difundida preferencia por el uso de la lógica matemática para entender el funcionamiento del sistema nervioso y del razonamiento humano;

— la instauración de la "meta" disciplina de la teoría de sistemas, que procura formular los principios generales que regirían todos los sistemas complejos; este abstracto enfoque comparativo ha dejado su impronta en muchas ramas de la ciencia, tales como la ingeniería (análisis de sistemas, teoría de los controles), biología (fisiología regulatoria, ecología), ciencias sociales (terapia familiar, antropología estructural, administración de empresas, urbanismo) y economía (teoría de los juegos);

— la teoría de la información como una

teoría estadística de la señal y de los canales de comunicaciones, que aún constituye la base de muchas tecnologías de la comunicación;

—los primeros ejemplos de robots parcialmente autónomos y sistemas autoorganizativos. Esta investigación constituía el centro del debate acerca de si la lógica (como antes se mencionó) bastaba para comprender el cerebro, pues pasaba por alto sus cualidades distribuidas y analógicas. Se plantearon otros modelos y teorías que gozaron de una recepción entusiasta. No obstante, este entusiasmo se extinguió pronto y sólo resurgió en la década de 1980, cuando revivió para constituir un importante enfoque de las CTC (tercera etapa). Para comprender el porqué de esta situación, primero debemos examinar los años cognitivistas; regresaremos en varias ocasiones a este importante tópico.

La lista es impresionante: tendemos a considerar que muchas de estas ideas y herramientas forman parte integral de nuestra vida. Aun así, ninguna de ellas existía antes de estos años de formación. Desde luego, esta creativa década dio muchos más frutos; sólo he dado una visión a vista de pájaro. En 1956 se dispersaron los principales actores de la fase cibernética, cuya unidad y vitalidad habían sido tan decisivas; muchos morirían poco después, y otros deberían convertirse en portadores de la llama de esa idea según la cual la mente era un mecanismo.

3

Segunda etapa. Los símbolos: la hipótesis cognitivista

Los cognitivistas entran en escena

Si en la década de 1940 nació la fase cibernética, se puede decir que 1956 fue el año que alumbró la segunda fase de las CTC. Durante este año, en dos encuentros celebrados en Cambridge y Dartmouth, nuevas voces como las de Herbert Simon, Noam Chomsky y Marvin Minsky manifestaron ideas que definirían los ejes principales de la moderna ciencia cognitiva.[9]

La intuición central que prevalecería en esas conferencias era la de que la inteligencia (incluida la inteligencia humana) se parece tanto a un ordenador o computador, en sus características esenciales, que la cognición se puede *definir* como la computación de representaciones simbólicas. Sin duda esta orientación no pudo surgir sin los cimientos echados durante la primera

[9] Para este período véase H. Gardner, *op. cit.*, capítulo 5.

etapa, que entre otras cosas creó la noción misma de computación. Lo que antes era una orientación tentativa —la mente como una forma lógica y, por ende, con una conducta similar a la de un ordenador— se eleva aquí a una *hipótesis* reconocida y se desean trazar límites respecto de sus raíces más amplias, exploratorias y multidisciplinarias, en las que la epistemología y las ciencias sociales y biológicas ocupaban un lugar preeminente con su múltiple complejidad. *Cognitivismo*[10] es una etiqueta conveniente para esta amplia pero bien perfilada orientación, que ha motivado muchos desarrollos científicos y tecnológicos desde 1956, en las áreas de la psicología cognitiva, la lingüística, buena parte de las neurociencias y, desde luego, la inteligencia artificial. Otras denominaciones que a veces se utilizan son: computacionismo (propiciada por Jerry Fodor) o procesamiento simbólico.

Síntesis de la doctrina

¿Qué significa decir que la cognición se puede definir como computación? Un cómputo es

[10] Esta denominación está justificada en J. Haugland (comp.), *Mind Design* (MIT Press, 1981). Para este capítulo he sacado gran provecho de la conferencia de D. Andler "Cognitivism — Orthodox and Otherwise. A New Phase?", perteneciente al ciclo *Man in the Age of Technology*, Atenas, junio de 1984.

una operación realizada mediante *símbolos*, es decir, mediante elementos que *representan* algo. Aquí la idea clave es la de representación o "intencionalidad", el término filosófico que designa algo que es "acerca de algo".[11] El argumento cognitivista es que la conducta inteligente supone la capacidad para representar el mundo de ciertas maneras. Así que no podemos explicar la conducta cognitiva a menos que demos por sentado que un agente actúa representando rasgos relevantes de las situaciones en que se halla. En la medida en que su representación de una situación sea exacta, la conducta del agente tendrá éxito (siempre que todas las demás cosas sean iguales).

Esta noción de la representación es —al menos desde el ocaso del conductismo— relativamente poco controvertida. Lo controvertido es el paso siguiente, en el cual el cognitivista afirma que el único modo de dar cuenta de la inteligencia y la intencionalidad consiste en formular la hipótesis de que la cognición consiste en actuar sobre la base de representaciones que *adquieren realidad física con la forma de un código simbólico en el cerebro o en una máquina.*

Según el cognitivista, el problema a resolver es el siguiente: cómo correlacionar la atribución de estados intencionales o representacionales

[11] Acerca de este tema, véase J. Searle, *Intentionality* (Cambridge University Press, 1983).

(como creencias, deseos e intenciones) con los cambios físicos que el agente sufre cuando actúa. En otras palabras, si deseamos afirmar que los estados intencionales tienen propiedades *causales*, tenemos que denostrar no sólo cómo esos estados son físicamente posibles, sino cómo pueden causar conducta. Aquí es donde entra la idea de *computación simbólica*: los símbolos tienen una realidad física y semántica, y esta realidad semántica condiciona la computación. En otras palabras, la computación es fundamentalmente semántica o representacional: la idea de computación no tiene sentido (en contraste con una operación simbólica aleatoria o arbitraria) sin tener en cuenta las relaciones semánticas existentes entre las expresiones simbólicas. (Este es el sentido del popular slogan "no hay computación sin representación"). Sin embargo, un ordenador digital opera sólo sobre la forma física de los símbolos que computa; no tiene acceso a su valor semántico. No obstante, sus operaciones están limitadas semánticamente porque los programadores han codificado toda distinción semántica relevante para su programa y la han expresado en la *sintaxis* de su lenguaje simbólico. En un ordenador la sintaxis refleja la proyección semántica o es paralela a ella. Los cognitivistas afirman, pues, que este paralelismo demuestra la realidad física y mecánica de la inteligencia y la intencionalidad (semántica). La hipótesis es pues que los ordenadores brindan un modelo mecánico del pensamiento o, en otras palabras,

que el pensamiento consiste en la computación física de símbolos. Las ciencias cognitivas se convierten en el estudio de esos sistemas cognitivos constitutidos por símbolos físicos.[12]

Para comprender adecuadamente esta hipótesis, es fundamental advertir en qué *nivel* se propone. El cognitivista no afirma que si abriéramos la cabeza de alguien y le miráramos el cerebro hallaríamos allí pequeños símbolos. El nivel simbólico tiene una dimensión física pero no es reductible al nivel físico. (Ello resulta evidente cuando recordamos que el mismo símbolo puede adoptar muchas formas físicas.) En consecuencia, una expresión simbólica poseedora de una realidad física se puede corresponder con un patrón global y altamente distribuido de actividad cerebral. Regresaremos luego sobre esta idea. Pero por ahora deseamos enfatizar que, además de los niveles de la física y la neurobiología, el cognitivismo postula un claro e irreductible nivel simbólico para explicar la cognición. Más aun, como los símbolos son ítems semánticos, los cognitivistas también postulan un tercer nivel semántico o representacional. (La irreductibilidad

[12] Véase Alan Newell, "Physical Symbol Systems", *Cognitive Science* 4:2 (1980), págs. 135-183; Herbert Simon, "Computer Science as Empirical Inquiry: Symbols and Search", reeditado en John Haugeland (comp.), *Mind Design* (Montgomery, Vermont, Bradford Books, 1981); y Zenon Pylyshyn, *Computation and Cognition: Toward a Foundation for Cognitive Science* (Cambridge, Massachusetts: Bradford Books/MIT Press, 1984).

de este nivel también resulta evidente cuando recordamos que el mismo valor semántico puede adoptar muchas formas simbólicas.)[13]

Esta multiplicidad de niveles en la explicación científica es muy reciente y constituye una de las principales innovaciones de las ciencias cognitivas. Los orígenes y la formulación inicial de la innovación como idea científica amplia nacieron en la era cibernética, pero los cognitivistas han hecho grandes aportes para una articulación filosófica más rigurosa.[14] Nos agradaría que el lector tuviera presente esta idea, pues cobrará nueva significación cuando en la siguiente etapa comentemos una idea emparentada con ella, la controvertida noción de *emergencia*.

El lector también debe advertir que la hipótesis cognitivista pone gran énfasis en las relaciones entre sintaxis y semántica. Como hemos mencionado, en un programa de computación la

[13] La irreductibilidad del nivel semántico es objeto de varias controversias entre los cognitivistas. Véase Stephen Stich, *From Folk Psychology to Cognitive Science: The Case Against Belief* (Cambridge, Massachusetts: Bradford Books/MIT Press, 1983); y Jerry Fodor, *Psychosemantics: The Problem of Meaning in the Philosophy of Mind* (Cambridge, Massachusetts: Bradford Books/MIT Press, 1987).

[14] Véase Jerry Fodor, "Special Sciences; or the Disunity of Science Considered as a Working Hypothesis," and "Computation and Reduction," reeditados en sus *Representations: Philosophical Essays on the Foundations of Cognitive Science* (Cambridge, Massachusetts: Bradford Books/MMIT Press, 1981).

sintaxis del código simbólico refleja su semántica. Ahora bien, en el caso del lenguaje humano, dista de ser obvio que todos los distingos semánticos relevantes para una explicación de la conducta estén reflejados sintácticamente. En verdad, se pueden esgrimir muchos argumentos filosóficos contra esta idea.[15] Más aun, aunque sabemos de dónde procede el nivel semántico de los cómputos de un ordenador (los programadores), no sabemos cómo obtienen su sentido las expresiones simbólicas que el cognitivista supone registradas en el cerebro.

El programa de investigación cognitivista se puede sintetizar en las respuestas a las siguientes preguntas:

Pregunta 1: ¿Qué es la cognición?

Respuesta: Procesamiento de información: manipulación de símbolos basada en reglas.

[15] Para una argumentación desde dentro de la filosofía analítica, veáse Hilary Putnam, "Computational Psychology and Interpretation Theory", reeditado en su *Realism and Reason: Philosophical Papers*, Volume 3 (Cambridge University Press, 1983). Para una crítica enactivista de esta idea, véase Terry Winograd y Fernando Flores, *Understanding Computers and Cognition: A New Foundation for Design* (Nueva Jersey: Ablex Press, 1986). Este problema también constituye el fundamento del ingenioso y hoy célebre experimento mental del "Cuarto chino", de John Searle, en su "Minds, Brains, and Programs", reeditado en John Haugeland (comp.), *Mind Design* (Montgomery, Vt.: Bradford Books, 1981).

Pregunta 2: ¿Cómo funciona?
Respuesta: A través de cualquier dispositivo que pueda representar y manipular elementos físicos discretos: los símbolos. El sistema interactúa sólo con la forma de los símbolos (sus atributos físicos), no su significado.

Pregunta 3: ¿Cómo saber que un sistema cognitivo funciona adecuadamente?
Respuesta: Cuando los símbolos representan apropiadamente un aspecto del mundo real, y el procesamiento de la información conduce a una feliz solución del problema planteado al sistema.

Obviamente el programa cognitivista que acabo de delinear no nació completo, como Atenea de la cabeza de Zeus. Lo presentamos con los beneficios de treinta años de perspectiva. Sin embargo, no sólo este audaz programa se ha consolidado plenamente, sino que ahora se lo *identifica* con las ciencias cognitivas. Pocos de sus participantes activos, y mucho menos el público en general, tienen en cuenta sus raíces o sus actuales desafíos y posibilidades. "El cerebro procesa información del mundo exterior" es una frase cotidiana que todos entienden. Decir que tal enunciado puede ser desorientador suena extraño, y de inmediato se tildará de "filosófica" a la subsiguiente conversación. Esta *ceguera* que el paradigma cognitivista ha introducido en el sentido común contemporáneo es capaz de poner

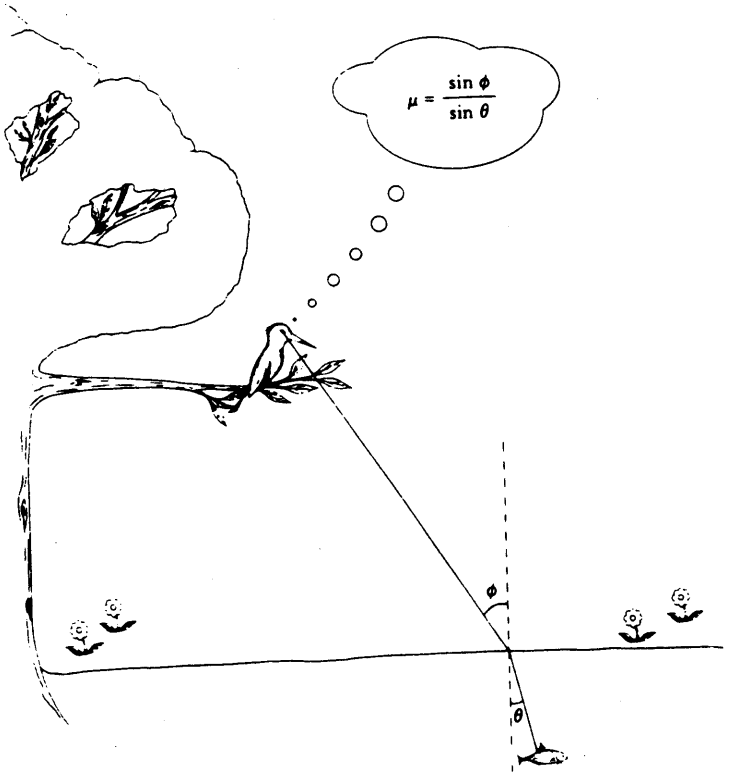


Figura 3. Una caricatura de la revista *Punch* que ilustra sucintamente la hipótesis cognitivista. Para capturar su presa, un martín pescador debe tener en el cerebro la representación de la ley de refracción de Snell.

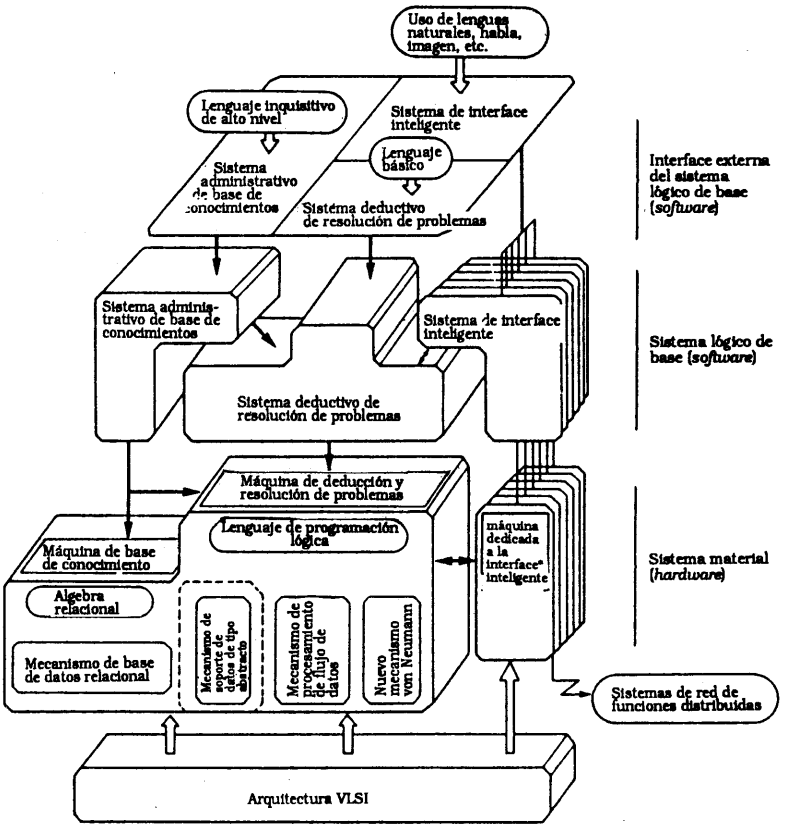
en jaque horizontes más amplios para el futuro de las CTC.

El vástago del cognitivismo: la inteligencia artificial

En ninguna parte son tan visibles las manifestaciones del cognitivismo como en la inteligencia artificial (IA), que es la *proyección literal* de la hipótesis cognitivista. Con el transcurso de los años, se han realizado muchos progresos teóricos y aplicaciones tecnológicas interesantes dentro de esta orientación: sistemas expertos, robótica, procesamiento de imágenes. Estos resultados han gozado de amplia difusión en publicaciones de divulgación, y no nos detendremos aquí en ellos.[16]

La IA y su base cognitivista alcanzaron su culminación social en el Programa de Quinta Generación ICOT del Japón. Por primera vez un plan concebido en escala nacional, concertando los esfuerzos de la industria, el gobierno y las universidades, ha convertido al Japón en líder en IA. El núcleo de este programa —el "cohete que se mandará a la Luna" en 1992— es un ordenador de quinta generación, un dispositivo cognitivo capaz de comprender el lenguaje humano y de escribir sus propios programas cuando el usua-

[16] Véase por ejemplo el número especial de *La Recherche* dedicado a *L'Intelligence artificielle* (Paris, Editions du Seuil, 1985).



* *Interface*: Dispositivos de comunicación hombre/máquina en tiempo real.

Figura 4. Esquema de un sistema informático de quinta generación, contemporánea punta de lanza del enfoque cognitivista de las CTC. (Fuente: Japan Information Processing Development Center.)

rio inexperto le propone una tarea. El corazón del programa ICOT utiliza PROLOG, un lenguaje de programación de alto nivel basado en la lógica de predicados, para la resolución de problemas y las *interfaces* entre patrones de representación. El programa ICOT ha causado reacciones inmediatas en Europa (el programa Esprit) y en los Estados Unidos. Es indudable que se trata de un decisivo campo de batalla comercial y técnico para la tecnología de la información (TI). Sin embargo, aquí no nos interesa si el cohete se construirá o no, sino que en efecto apunte hacia la Luna. Volveremos luego sobre esto.

Las ciencias cognitivas

La hipótesis cognitivista tiene en la IA su proyección más literal. Su propósito complementario consiste en el estudio de sistemas cognitivos naturales y biológicos, muy especialmente el hombre. También aquí la principal herramienta explicativa ha sido la formulación computacional de la representación. Las representaciones mentales son asimiladas a elementos de un sistema formal a los que la actividad de la mente da su matiz interpretativo: creencias, deseos, planes y demás. Al contrario de lo que ocurre en la IA, pues, aquí encontramos un interés en los sistemas cognitivos *naturales*, y se da por sentado que sus representaciones cognitivas son acerca de algo para el sistema (es decir, son intencionales).

Por ejemplo, se presenta a los sujetos figuras geométricas y se les pide que las hagan rotar mentalmente. Los informes insisten en que la dificultad de la tarea, medida según el tiempo que se tarda en realizarla, depende del número de grados de libertad en que se debe hacer rotar la figura: en un espacio plano o tridimensional. Es decir, da la impresión de que tuviéramos una "pantalla" mental interna en donde las figuras rotan como en una pantalla de televisión.[17] Con el tiempo estos experimentos condujeron a una teoría explícita que postulaba reglas de operación del espacio mental similares a los que usan los ordenadores para desplegar los datos almacenados. Los investigadores sugerían que existe una interacción entre las operaciones de tipo lingüístico y las operaciones de tipo visual, y que juntas generan nuestro ojo interno.[18] Este enfoque ha generado una abundante literatura a favor y en contra, y cada nivel de las observaciones ha recibido diversas interpretaciones. Los cognitivistas acérrimos, tales como Zenon Pylyshyn, argumentan que las imágenes son simplemente epifenómenos subjetivos de una computación simbólica más profunda.[19] Sin embargo, el estudio de las imágenes es un ejemplo representativo del modo en que actúa el enfoque cognitivista cuando estudia los fenómenos mentales.

[17] R. Shepard y J. Metzler, *Science* 171:701-3, 1971.

[18] S. Kosslyn, *Psychol. Rev.* 88:46-66, 1981.

[19] Véase *Beh. Brain Sci.* 2:535-81, 1979, y H. Gardner, *The Mind's New Science, op. cit.*

El procesamiento de la información en el cerebro

El cognitivismo también ha influido muchísimo en los estudios del cerebro. Aunque téoricamente el nivel simbólico del cognitivismo es compatible con muchos enfoques sobre el cerebro, en la práctica casi toda la neurobiología (y su enorme masa de pruebas empíricas) está imbuida de la perspectiva cognitivista del procesamiento de la información. Así comienza un popular libro de texto sobre la neurociencia: "El cerebro es un activo conjunto de células que continuamente recibe, elabora y percibe información, y toma decisiones".[20] A menudo ni siquiera se cuestionan los orígenes y los supuestos de esta perspectiva.

Un ejemplo sobresaliente son las dos décadas de estudios sobre la corteza visual, una zona del cerebro en la que es fácil detectar respuestas eléctricas de las neuronas cuando se presenta al animal una imagen visual. Se declaró tempranamente que era posible clasificar las neuronas corticales como *detectores de rasgos* que responden a ciertos atributos del objeto de marras: su orientación, contraste, velocidad, color y demás. En concordancia con la hipótesis cognitivista, se suele considerar que estos resultados dan respaldo biológico a la idea de que el cerebro recoge

[20] S. Kuffler y J. Nichols, *From Neuron to Brain* (Boston: Sinauer Associates, 1975), pág. 3.

información visual a partir de la retina por intermedio de las neuronas detectoras de rasgos de la corteza, y que luego la información pasa a posteriores etapas del cerebro para nuevos procesamientos (categorización conceptual, asociaciones de memoria, y eventualmente la acción).[21]

En su forma más extrema, esta visión del cerebro está expresada en la doctrina de la "célula abuela" de Barlow, en la que hay una correspondencia entre conceptos (tal como el concepto que alguien tiene de su abuela) o percepciones y neuronas específicas,[22] un equivalente IA de los sensores electrónicos y las etiquetas acanaladas. Esta posición extrema está perdiendo popularidad,[23] pero la idea básica de que el cerebro es un dispositivo para procesar información y que reacciona selectivamente ante ciertas características ambientales persiste en el núcleo de la neurociencia moderna y de la percepción del público. Luego volveremos a hablar de este problema.

[21] El libro de P. Buser y M. Imbert, *Vision* (París: Hermann, 1986), ilustra bien esta tendencia en Francia.

[22] H. Barlow, "Single Units and Sensation: A Neuron Doctrine for Perceptual Psychology", *Perception* 1 (1972): 371-394.

[23] Véase por ejemplo, la crítica a Barlow emprendida por David Marr en su trabajo *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information* (Nueva York: W. H. Freeman and Company, 1982).

Síntesis del disenso

El cognitivismo en las CTC constituye sin duda un programa de investigación sólido y bien definido que incluye instituciones prestigiosas, publicaciones especializadas, tecnología aplicada e intereses comerciales internacionales. La mayoría de las personas especializadas en IA (y TI) suscribirían —a sabiendas o no— al cognitivismo. A fin de cuentas, si uno se dedica cotidianamente a usar programas LISP o a hallar neuronas para tareas bien definidas, no podría ser de otra manera. En este ensayo deseamos llamar la atención sobre la profundidad de este *compromiso social* de un amplio sector de la comunidad de investigadores de CTC. Nos interesa examinar los fundamentos de las CTC cognitivistas para aclarar las bases del disenso. El disenso esencial respecto de los puntos de vista establecidos en las CTC de hoy cobra dos formas básicas:

—La crítica de la computación simbólica en cuanto portadora adecuada de las representaciones.

—La crítica de la pertinencia de la noción de representación en cuanto componente fundamental de las CTC.

En los dos capítulos siguientes analizaremos ambas líneas de disenso.

4

Tercera etapa. La emergencia: una alternativa ante la orientación simbólica

La autoorganización: las raíces de una alternativa

Ya en los primeros años de la cibernética se propusieron otras posibilidades ante el aplastante dominio de la lógica como enfoque predominante en las ciencias cognitivas. En las Conferencias Macy,[24] por ejemplo, se comentó a menudo que en los cerebros reales no hay reglas ni un procesador lógico central, y que la información no está almacenada en lugares precisos. En cambio, era evidente que el cerebro operaba a partir de interconexiones masivas, de forma distribuida, de modo que las conexiones entre conjuntos de neuronas cambian como resultado de la experiencia. En síntesis, estos conjuntos presentan una capacidad autoorganizativa que no es propia de la lógica. En 1958 Frank Rosenblatt

24 Para fuentes sobre esos años formativos, véase nota 2.

construyó el "perceptrón", un aparato simple con cierta capacidad de reconocimiento, basándose en los cambios de conectividad entre componentes semejantes a neuronas[25]; análogamente, W. R. Ashby realizó el primer estudio de la dinámica de sistemas muy grandes con interconexiones aleatorias, mostrando que exhiben conductas globales coherentes.[26] Aclararemos y ejemplificaremos estas ideas en las páginas siguientes.

La historia quiso que estos otros enfoques fueran literalmente borrados del escenario intelectual en beneficio de las ideas computacionales comentadas en el Capítulo 3. Sólo a fines de la década de 1970 estas ideas revivieron explosivamente, al cabo de veinticinco años de predominio de la ortodoxia cognitivista (lo que Daniel Dennett ha denominado irónicamente "computacionalismo de la Alta Iglesia").[27] Por cierto uno de los factores que contribuyó a este renovado interés fue el redescubrimiento paralelo de las

[25] Frank Rosenblatt, *Principles of Neurodynamics: Perceptrons and the Theory of Brain Dynamics* (Spartan Books, 1962).

[26] Para más datos sobre los complejos orígenes de las ideas acerca de la autoorganización, véase Isabelle Stengers, "Les généalogies de l'auto-organisation", *Cahier du CREA* (París) Nº 8, págs. 7-105..

[27] Daniel C. Dennett, "Computer Models and the Mind — A View from the East Pole", *Times Literary Supplement*, 14 de diciembre de 1984. También reeditado como "The Logical Geography of Computational Approaches: A view from the East Pole", en M. Brand y M. Harnish (comps.), *The Representation of Knowledge* (Tucson: University of Arizona

ideas autoorganizativas en física y en matemática no lineal,[28] así como el fácil acceso a ordenadores rápidos, como veremos a continuación.

La nueva motivación para echar un segundo vistazo a la autoorganización se basaba en dos reconocidas lagunas del cognitivismo. La primera es que el procesamiento de información simbólica se basa en reglas *secuenciales*, aplicadas una por vez. Este "cuello de botella von Neumann" se convierte en una seria limitación cuando la tarea requiere gran cantidad de operaciones secuenciales (tales como el análisis de imágenes o el pronóstico meteorológico). La continua búsqueda de algoritmos de procesamiento *paralelo* ha tenido poco éxito porque toda la ortodoxia computacional parece ir precisamente contra esa corriente.

Una segunda limitación importante es que el procesamiento simbólico está localizado: la pérdida o disfunción de cualquier parte de los símbolos o reglas del sistema deriva en un grave daño. En cambio, una operación *distribuida* resulta muy deseable, para que al menos haya una relativa equipotencialidad e inmunidad ante las mutilaciones.

Press, 1986). Para una visión diferente de estas cuestiones históricas, véase también Marvin Minsky y Seymour Papert, *Perceptrons*, prólogo y epílogo a la visión revisada de 1987 (Cambridge, Massachusetts: MIT Press, 1969/1987).

[28] El libro de I. Prigogine y I. Stengers, *La Nouvelle alliance* (París: Gallimard, 1981), constituye un buen ejemplo de esta tendencia, que aquí no comentamos en absoluto.

Los resultados de la experiencia de las dos primeras décadas de predominio cognitivista se expresan claramente señalando una convicción que gradualmente cobró arraigo en la comunidad de investigadores: es preciso invertir los papeles del experto y del niño en la escala de desempeños. Los primeros intentos se proponían resolver los problemas más generales, tales como la traducción del lenguaje natural o el diseño de soluciones generales para los problemas. Se pensaba que estos intentos, que trataban de imitar la pericia de los expertos, abordaban los problemas difíciles e interesantes. A medida que los intentos se volvían más modestos y localizados, fue evidente que la inteligencia más profunda y fundamental es la del bebé que puede adquirir el lenguaje a partir de manifestaciones diarias y dispersas, y que puede distinguir objetos significativos a partir de lo que parece ser un mar de luces. Las arquitecturas cognitivistas se habían alejado en exceso de las inspiraciones biológicas. No se trata de reducir lo cognitivo a lo biológico, sino de que las tareas más comunes, aun las emprendidas por diminutos insectos, se realizan más deprisa que cuando se intentan con una estrategia computacional como la que propone la ortodoxia cognitivista. Análogamente, los neurobiólogos dan por sentada la plasticidad del cerebro ante las lesiones, o la flexibilidad de la cognición biológica para adaptarse a nuevos ámbitos sin perder toda su competencia, pero ella no existe en el paradigma computacional en cuanto tal.

La estrategia conexionista

Así el cerebro ha vuelto a convertirse en fuente de metáforas e ideas para otros campos de las ciencias cognitivistas en esta orientación alternativa. Aquí no partimos de descripciones simbólicas abstractas, sino de una hueste de componentes de tipo neural, no inteligentes, que, apropiadamente conectados, presentan interesantes propiedades *globales*. Estas propiedades globales corresponden a las aptitudes cognitivas que se estudian.

Todo el enfoque depende, pues, de introducir las conexiones apropiadas, lo cual se suele hacer mediante una regla para el gradual *cambio* de conexiones a partir de un estado inicial bastante arbitrario. La regla de aprendizaje mejor explorada es la "Regla de Hebb". En 1949 Donald Hebb sugirió que el aprendizaje se podía basar en cambios cerebrales que surgen del grado de actividad *correlacionada* entre las neuronas: si dos neuronas tienden a actuar en conjunto, su conexión se refuerza; de lo contrario disminuye. Por lo tanto, la conectividad del sistema se vuelve inseparable de su historia de transformación y se relaciona con la clase de tarea que se propone al sistema. Como la verdadera acción se produce en el nivel de las conexiones, se ha propuesto el nombre de *conexionismo* (a menudo llamado neoconexionismo, para distinguirlo del paleoconexionismo de los

años cibernéticos) para esta línea de investigación.[29]

Un factor decisivo en el explosivo interés que despierta este enfoque fue la introducción de métodos eficaces para seguir los cambios que se producen en estas redes. Se ha dado gran atención a la introducción de medidas estadísticas que brindan al sistema una función de "energía" global que nos permite seguir cómo llega el sistema a estados convergentes.[30]

Veamos un ejemplo, ilustrado en la Figura 5.1. Tomemos un número total N de elementos simples semejantes a neuronas y conectémoslos entre sí. Luego presentemos a este sistema una sucesión de patrones tratando algunos de sus nódulos como extremos sensoriales (digamos una retina). Al cabo de cada presentación, dejemos que el sistema se reorganice reacomodando sus conexiones de acuerdo con un principio heb-

[29] El nombre fue propuesto por J. Feldman y D. Ballard, "Connectionist models and their properties", *Cognitive Science* 6: 1982, págs. 205-254. Para un extenso comentario acerca de los modales actuales de esta tendencia, véase J. McClelland y D. Rummelhart (comps.), *Parallel Distributed Processing: Studies on the Microstructure of Cognition* (Cambridge, Massachusetts: MIT Press, 1986).

[30] La principal idea se debe aquí a John J. Hopfield, "Neural Networks and Physical Systems with Emergent Computational Abilities", *Proceedings of the National Academy of Sciences* (USA), 79 (1982):2554-558. Véase también David W. Tank y John J. Hopfield, "Collective Computation in Neuronlike Circuits", *Scientific American*, diciembre de 1987.

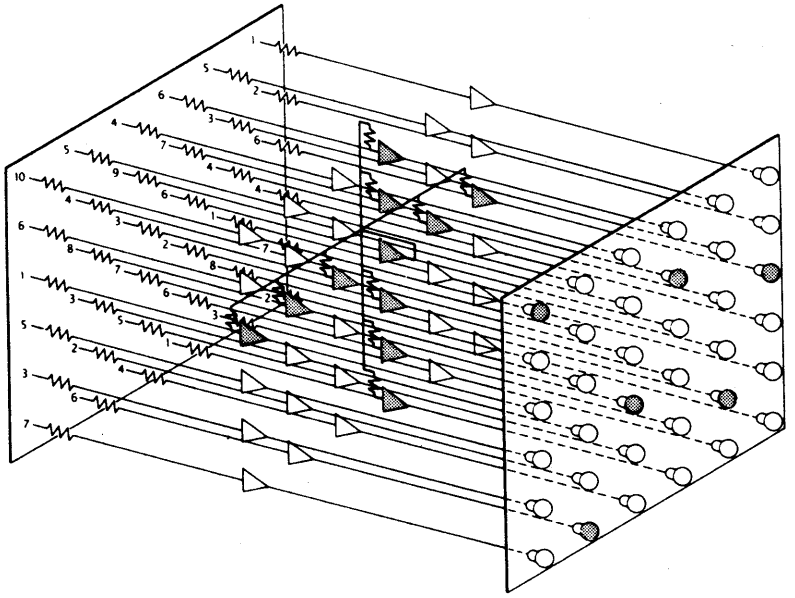


Figura 5. Red de Hopfield. Un circuito de optimización para resolver el problema de asignación de tareas. Consiste en una red de *n-flops* interconectadas. Los amplificadores de cada hilera y columna están enlazados por conexiones inhibitorias que imponen la restricción de que sólo un amplificador por hilera o columna pueda hallarse en estado +1. Como cada uno de los 36 amplificadores de esta red inhibe a otros 10 amplificadores, hay 360 conexiones en total. El diagrama presenta las conexiones de uno de los amplificadores. Los amplificadores reciben entradas de corriente proporcionales a las tasas de reducción. Los amplificadores que corresponden a la mejor solución —la combinación de entradas [*inputs*] que suman el total mayor— emiten un + 1 y el resto emite un 0. Las emisiones [*outputs*] pueden controlar un panel, en este caso un conjunto de bombillas eléctricas. (Fuente: Hopfield y Tank, "Collective Computation in Neuronlike Circuits", *Scientific American*, diciembre 1987)

biano, es decir, incrementando los enlaces entre las neuronas que actúan juntas ante el ítem presentado. La presentación de la lista de patrones constituye la fase de aprendizaje del sistema.

Después de la fase de aprendizaje, cuando volvemos a presentar uno de estos patrones al sistema, éste lo reconoce *en el sentido* de que adopta un estado global singular o configuración interna que presuntamente reproduce el ítem aprendido. Este reconocimiento es posible siempre que la cantidad de patrones presentados no sea superior a cierta fracción del número total de neuronas que participan (alrededor de 0,15 N). Más aun, el sistema realiza un correcto reconocimiento aunque se le presente el patrón con ruido añadido, o aunque el sistema esté parcialmente mutilado.[31]

Emergencia y autoorganización

Este ejemplo es sólo uno de toda una clase de redes neurales o modelos conexionistas, tal como veremos más adelante. Pero antes de ello es preciso ampliar nuestro comentario para ver

[31] Hay muchas variantes sobre estas ideas. Véase G. Hinton, T. Sejnowsky, y D. Ackley, "A Learning Algorithm for Boltzman Machines", *Cognitive Science* 9 (1985): 147-169; y G. Toulouse, S. Dehaene y J. Changeux, *Proceedings of the National Academy of Sciences (USA)* 83, 1986: 1695-1698.

qué está en juego en el estudio de estas redes. La estrategia, como dijimos, consiste en construir un sistema cognitivo no a partir de símbolos y reglas, sino de componentes simples que se conectarían dinámicamente entre sí de maneras densas. En este enfoque, cada componente opera sólo en su ámbito *local*, de modo que no hay un agente externo que, por así decirlo, haga girar la manivela del sistema. Pero, dada la constitución de la red del sistema, hay una cooperación *global* que *emerge* espontáneamente cuando todas las "neuronas" participantes alcanzan un estado mutuamente satisfactorio. En tal sistema, pues, no se requiere una unidad procesadora central que guíe toda la operación. Este tránsito de las reglas locales a la coherencia global es el corazón de lo que en los años cibernéticos se denominaba autoorganización.[32] Hoy la gente prefiere hablar de propiedades emergentes o globales, dinámica de red, redes no lineales, sistemas complejos o aun de sinergia.[33]

No hay una teoría formal unificada de las propiedades emergentes. No obstante, resulta claro que se han hallado propiedades emergentes en todos los dominios: vórtices y láseres, oscila-

32 Véase, por ejemplo, Heinz von Foerster (comp.), *Principles of Self-Organization* (Nueva York: Pergamon Press, 1962).

33 Para una extensa discusión sobre este punto, véase P. Dumouchel y J. P. Dupyy (comps.), *L'Auto-Organisation: De la Physique au Politique* (París: Editions du Seuil, 1983).

ciones químicas, redes genéticas, patrones de desarrollo, genética de población, redes de inmunidad, ecología y geofísica. Lo que tienen en común estos diversos fenómenos es que en cada caso una red permite surgir nuevas propiedades, las cuales los investigadores procuran entender en toda su generalidad.[34] El concepto de "atractor", procedente de la teoría de los sistemas dinámicos, es muy útil para captar las propiedades emergentes que estos diversos sistemas tienen en común. Como esta idea será importante para el resto de nuestro análisis, detengámonos a examinar un ejemplo.[35]

Pensemos en un "autómata celular", una unidad simple que recibe información de dos vecinos inmediatos y comunica su estado interno a dichos vecinos. Supongamos que la célula o unidad puede estar en sólo dos estados (0 y 1, activo o inactivo) y que la regla que gobierna el cambio en cada autómata es simplemente una función (de Boole) de dos argumentos (tales como

[34] En los Estados Unidos el Santa Fe Institute for the Study of Complex Systems, y la creación de una nueva publicación, *Complex Systems*, son claros síntomas de esta tendencia creciente. Remito al lector ávido de detalles a dichas fuentes.

[35] Una introducción accesible a la moderna teoría de los sistemas dinámicos es R. Abraham y C. Shaw, *Dynamics: The Geometry of Behavior* (Santa Cruz: Aerial Press, 3 vols. 1985). Para introducciones menos técnicas, véase también James Crutchfield y otros autores, "Chaos", *Scientific American*, diciembre de 1986; y James Gleick, *Chaos: the Making of a New Science* (Nueva York: Viking Press, 1987).

"y" u "o excluyente"). Como podemos escoger tal función para cada uno de los dos estados en que está el autómata celular, la operación de cada unidad está completamente definida por un par de funciones de Boole (Figura 6-1).

En vez de trabajar con una red compleja, simplemente conectamos una serie de tales unidades elementales en disposición circular, para que no haya entrada ni salida de información en todo el círculo, sino sólo acciones internas. Para el propósito ilustrativo, sin embargo, resulta más fácil cortar el anillo y representarlo linealmente, con las células en estado 1 indicadas por un cuadrado negro, y el estado opuesto indicado por un espacio en blanco. Por lo tanto, en la ilustración de la Figura 6-2[36], la posición celular va de izquierda a derecha (módulo, la longitud del anillo).

El autómata celular anular adquiere cierta dinámica cuando, a partir de un estado aleatorio, dejamos que cada célula calcule su siguiente estado en cada nueva unidad temporal (discreta) de manera sincrónica (es decir, todas las células llegan juntas a sus respectivos estados). En el gráfico, representamos el instante inicial en la hilera superior, y los sucesivos instantes del tiempo hacia abajo. Así el estado sucesivo de la misma célula se puede leer como una columna, y

[36] Figura y simulaciones tomadas de F. Varela, *Autonomie et connaissance* (París, Éditions du Seuil, 1989), capítulo 11.

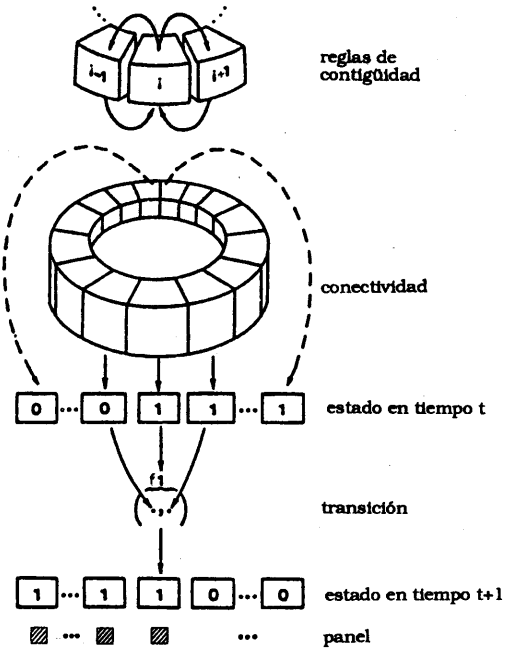


Figura 6.1. Para construir un sistema simple dotado de autoorganización, comencemos por una cadena de elementos simples similares a las neuronas de McCulloch-Pitts de la Figura 2. Conectémoslas de tal modo que cada neurona reciba influencias sólo de sus vecinas contiguas, e influya sobre ellas. La configuración se cierra sobre sí misma formando un anillo. La conducta de cada neurona se determina mediante la elección de una regla lógica. En cada momento del tiempo, el estado de todo el anillo se puede comprobar con sólo mostrar en el panel neuronas activas como un cuadrado negro, y las neuronas inactivas como blancos. Para seguir lo que sucede en el tiempo, el primer instante se muestra como la primera hilera, el segundo instante como la segunda hilera, y así sucesivamente. (Tomado de F. Varela, *Autonomie et connaissance, op. cit.*)

el estado simultáneo de todas las células se puede leer como una hilera. En todas las simulaciones presentadas en la Figura 6-2 el anillo estaba compuesto de 80 células, y su estado inicial fue escogido al azar.

Es interesante observar que aun esta red simple, casi mínima, posee una gran capacidad autoorganizativa. S. Wolfram ha realizado recientemente un examen de dicha capacidad.[37] No recapitularemos aquí su trabajo. Para nuestros propósitos basta con señalar que dinámicamente estos anillos se clasifican en cuatro clases principales o atractores, como se ilustra en la Figura 6-2. Una primera clase exhibe un solo atractor, que induce a todas las células a volverse homogéneamente activas o inactivas. Para una segunda y más interesante clase de anillos, las reglas suscitan periodicidades espaciales, es decir, algunas células permanecen activas mientras que otras no. Para una tercera clase, las reglas suscitan ciclos espaciotemporales de longitud dos o más. Estas dos últimas clases se corresponden con atractores cíclicos. Finalmente, para unas pocas reglas, la dinámica parece suscitar atractores caóticos, en los cuales no se detecta ninguna regularidad en el espacio ni en el tiempo.

El punto básico que ilustramos aquí es que

37 Véase S. Wolfram, "Statiscal mechanics of cellular automata", *Reviews of Modern Physics* 55 (1983), págs. 601-644; y "Cellular automata as models of complexity", *Nature* 311 (1984), pág. 419.

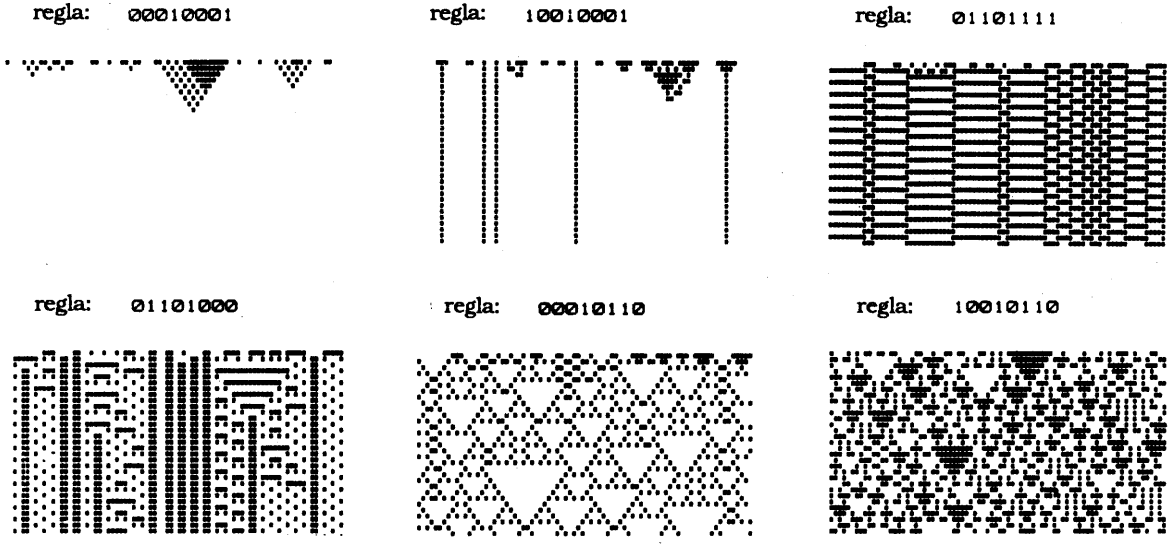


Figura 6.2. Al estudiar la conducta de los anillos con reglas diferentes, se observa que rápidamente surgen esquemas globales de tales sistemas, resultado de una actividad cooperadora entre todas las neuronas constitutivas. Se obtienen, entonces, representaciones visuales de atractores a partir de estos sistemas autoorganizados. (Según F. Varela, 1989, *op. cit.*)

la emergencia de patrones o configuraciones globales en sistemas de elementos interactuantes no es una rareza de casos aislados ni es exclusiva de los sistemas neurales. De hecho, parece difícil que ningún compuesto densamente conectado escape a las propiedades emergentes, así que las teorías de tales propiedades son un enlace natural para diversos niveles de descripciones en los fenómenos naturales y cognitivos. Teniendo en cuenta esta perspectiva más amplia de la autoorganización, regresemos ahora a las redes neurales y al conexionismo.

El conexionismo en la actualidad

Las teorías conexionistas brindan elegantes modelos funcionales para diversas e interesantes aptitudes cognitivas, tales como el reconocimiento rápido, la memoria asociativa y la generalización categórica, como vimos en el ejemplo de la Figura 5. El actual entusiasmo por esta orientación está justificado por diversas razones. Primero, la IA cognitivista y la neurociencia contaban con pocos resultados convincentes para explicar (o reconstruir) los procesos cognitivos que acabamos de presentar. Segundo, los modelos conexionistas están mucho más cerca de los sistemas biológicos, así que se puede trabajar con un grado de integración entre la IA y la neurociencia que hasta el momento era impensable. Por último, los modelos son tan generales como

para aplicarlos, con escasas modificaciones, a diversas áreas, tales como la visión o el reconocimiento del lenguaje.

Hay diversos ejemplos de estados neurales emergentes para tareas que no requieren aprendizaje, tales como los movimientos oculares o el desplazamiento balístico de los miembros. Obviamente, la mayoría de las tareas cognitivas que deseamos comprender involucran transformaciones dependientes de la experiencia, de allí el interés en aprender reglas como la de Hebb, que introdujimos en nuestro primer ejemplo. Tales reglas de aprendizaje brindan a una red neural no sólo configuraciones emergentes (como ocurría aun con nuestro simple autómata celular), sino la aptitud para sintetizar nuevas configuraciones de acuerdo con la experiencia.

No reseñaremos aquí este explosivo campo de investigación en las redes neurales plásticas y sus aplicaciones al estudio del cerebro y la inteligencia artificial.[38] Baste con destacar que actualmente se exploran dos clases principales de métodos de aprendizaje. El primero, ilustrado por la regla de Hebb e inspirado por los mecanismos cerebrales, es el aprendizaje por *correlación*: se presenta al sistema una serie de ejemplos que lo condicionan para futuros encuentros. El segun-

[38] Para una reseña reciente y representativa, véase *IEEE First International Conference on Neural Networks*, 4 vols. (IEEE Press, 1987).

do es el aprendizaje por imitación, es decir mediante un modelo que actúa como *instructor* activo. Esta estrategia es la propuesta por Rosenblatt en su Perceptrón. En su versión moderna se la conoce como "retropropagación". En esta técnica, los cambios en las conexiones neuronales del interior de la red (llamadas unidades ocultas) se asignan de tal modo de reducir al mínimo la diferencia entre la reacción de la red y lo que se espera de ella.[39] Aquí el aprendizaje se parece a alguien que trata de imitar a un instructor. NetTalk, un célebre y reciente ejemplo de este método, es un conversor grafemas-fonemas que opera a partir de ciertas páginas en inglés que se le presentan en su fase de aprendizaje. Como resultado, NetTalk puede leer en voz alta un nuevo texto en lo que muchos testigos consideran un inglés deficiente pero comprensible.[40]

Las emergencias neuronales

Los trabajos recientes han brindado pruebas detalladas de que las propiedades emergentes son fundamentales para la operación del cerebro. Ello

[39] La idea en su forma moderna es debida a D. Rummelhart, G. Hinton y R. Williams, en Rummelhart y McClelland, *Parallel Distributed Processing*, capítulo 8.

[40] Véase T. Sejnowski y C. Rosenbaum, "NetTalk: A parallel network that learns to read aloud", TR JHU/EECS-86, Johns Hopkins University.

no es sorprendente si observamos los detalles de la anatomía cerebral. De hecho, desde los tiempos de Sherrington y Pavlov, la comprensión de las propiedades globales distribuidas ha sido la Meca de la neurociencia, y no es fácil llegar a ella. Las razones de estas dificultades han sido tanto técnicas como conceptuales. Han sido técnicas porque no es fácil saber qué hacen simultáneamente millares de neuronas dispersas en todo el cerebro. Sólo recientemente algunos métodos se han vuelto realmente eficaces.[41] Pero las dificultades también han sido conceptuales, pues, como antes señalamos, en las décadas de 1960 y 1970 los neurocientíficos tuvieron una gran propensión a mirar el cerebro con gafas cognitivistas. Las metáforas relacionadas con el procesamiento de información, basadas en la creencia de que el cerebro se puede describir como un ordenador de von Neumann, estaban más en boga que las descripciones basadas en redes emergentes.

Sin embargo, las metáforas relacionadas con el procesamiento de información tienen una utilidad muy restringida. Por ejemplo, aunque las neuronas de la corteza visual manifiestan reacciones determinadas ante "rasgos" específicos de los estímulos visuales, tal como hemos descrito, estas reacciones sólo se producen en un animal

[41] Para una interesante compilación de recientes ejemplos y comentarios, véase G. Palm y A. Aersten (comps.), *Brain Theory* (Berlín: Springer-Verlag, 1986).

anestesiado cuyo ámbito (interno y externo) está muy simplificado. Cuando el animal estudiado está consciente y activo en un ámbito sensorial más normal, resulta cada vez más evidente que las respuestas neuronales estereotipadas que describimos antes se vuelven muy sensibles al contexto. Por ejemplo, hay claros efectos producidos por la inclinación corporal o la estimulación auditiva.[42] Más aun, las características de la respuesta neuronal dependen directamente de neuronas localizadas lejos de sus campos receptivos.[43] Aun un cambio de postura, ante un idéntico estímulo sensorial, altera las respuestas neuronales en la corteza visual primaria, demostrando que, a pesar de la aparente separación, el *motorium* está en consonancia con el *sensorium*.[44] Parece antinatural hacer una descripción simbólica, punto por punto, de un sistema configurado de este modo.

Por lo tanto, los neurocientíficos han comprendido que es preciso estudiar las neuronas como miembros de grandes conjuntos que aparecen y desaparecen constantemente a través de

42 Para los efectos de la inclinación corporal, véase G. Horn y R. Hill, *Nature* 221 (1974), págs. 185-187. Para los efectos de la estimulación auditiva, véase M. Fishman y C. Michael, *Vision Research* 13 (1973), pág. 1415; y F. Morell, *Nature* 238 (1972), págs. 44- 46.

43 Véase J. Allman, F. Meizen y E. McGuiness, *Annual Review of Neuroscience* 8 (1985), págs. 407-430.

44 M. Abeles, *Local Circuits* (Nueva York: Springer, 1984).

sus interacciones cooperativas, y donde cada neurona tiene respuestas múltiples y cambiantes en un modo que depende del contexto. Una regla para la constitución del cerebro es que si una región (núcleo, capa) A se conecta con B, luego B se conecta recíprocamente con A. Esta ley de reciprocidad tiene sólo dos o tres excepciones menores. El cerebro es así un sistema altamente cooperativo: las densas interconexiones entre sus componentes implican que eventualmente todo ocurrirá en función de *todos* esos componentes.

Esta clase de cooperación se sostiene tanto local como globalmente: funciona dentro de los subsistemas cerebrales y en las conexiones entre dichos subsistemas. Uno puede tomar el cerebro entero y dividirlo en subsecciones, según las clases de células y zonas, tales como el tálamo, el hipotálamo, los pliegues corticales, etc. Estas subsecciones están conformadas por complejas redes celulares, pero también se relacionan entre sí conformando una red. En consecuencia, todo el sistema cobra una coherencia interna en patrones intrincados, aunque no podamos decir exactamente cómo ocurre. Si uno moviliza artificialmente el sistema reticular, un organismo cambia de conducta: por ejemplo, pasa de estar despierto a estar dormido. Sin embargo, este cambio no indica que el sistema reticular controle la vigilia. Ese sistema es más bien una forma de arquitectura, dentro del cerebro, que *permite* la manifestación de ciertas coherencias internas. Pero cuando tales coherencias se manifiestan, no

se debe simplemente a un sistema particular. El sistema reticular es *necesario* pero no *suficiente* para ciertos estados coherentes, tales como la vigilia y el sueño. Lo que está dormido o despierto es el animal, no las neuronas reticulares.

Veamos lo que ocurre en la percepción visual en sus etapas periféricas. El primer diagrama de la Figura 5.5 despliega las sendas visuales cuando uno mira el cerebro desde abajo. El nervio óptico va desde los ojos hasta una región del tálamo llamada núcleo geniculado lateral (NGL) y desde allí a la corteza visual. La descripción estándar del procesamiento de información (todavía hallada en libros de texto y en publicaciones populares) es que la información entra por los ojos y es retransmitida secuencialmente a través del tálamo hasta la corteza, donde se llevan a cabo "más procesamientos". Pero si miramos atentamente la configuración de todo el sistema, encontramos pocos elementos que respalden esta idea de secuencialidad. El diagrama de la Figura 7 retrata el modo en que el NGL está encastrado en la red cerebral. Es èvidente que el 80 por ciento de lo que ve cualquier célula del NGL no proviene de la retina, sino de la densa interconectividad de otras regiones del cerebro. Más aun, son más las fibras que *bajan* desde la corteza al NGL que las que suben en dirección inversa. Considerar las sendas visuales como procesos secuenciales parece totalmente arbitrario; de la misma manera podríamos considerar que la secuencia se mueve en la dirección inversa.

De manera que aun en el extremo más periférico del sistema visual las influencias que el cerebro recibe del ojo se topan con más actividad que fluye desde la corteza. El encuentro de estos dos conjuntos de actividad neuronal es una etapa en la emergencia de una nueva configuración coherente, que depende del éxito del cotejo entre la

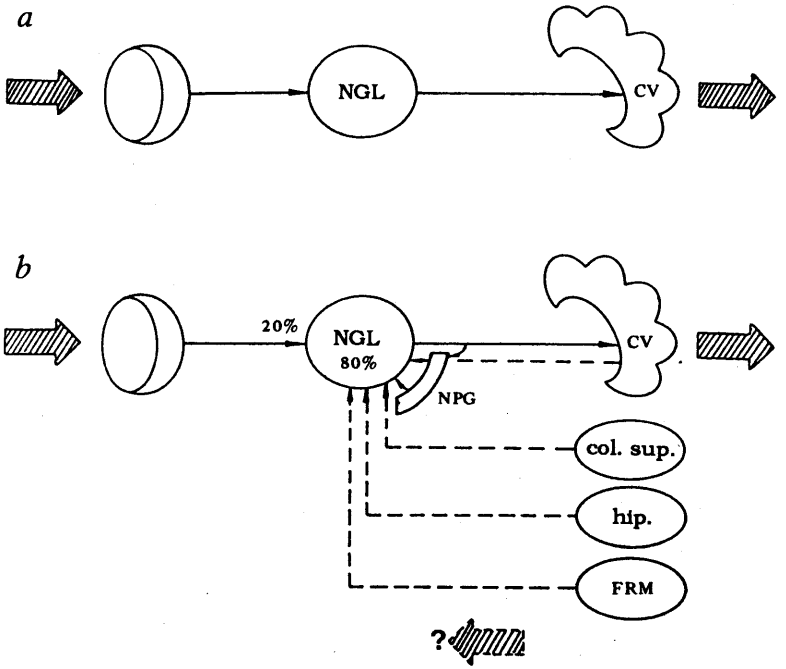


Figura 7. El diagrama de las conexiones del sistema visual de los mamíferos. NPG: núcleo perigeniculado; col. sup: colículo superior; hip.: hipotálamo; FRM formación reticular mediana; C.V.: corteza visual; NGL: núcleo geniculado lateral.

actividad sensorial y la conformación "interna" de la corteza primaria.[45] La corteza visual primaria es, no obstante, sólo uno de los socios de este circuito neuronal local específico en el nivel del NGL. Otros socios desempeñan papeles igualmente activos, tales como la formación reticular, las fibras procedentes del colículo superior, o la consecuente actividad de las neuronas que controlan los movimientos oculares.[46] La conducta de todo el sistema se parece más a una animada charla en una fiesta que a una cadena de mando.

Lo que hemos descrito para la NGL y la visión es válido para todo el cerebro. La visión es un ejemplo cómodo porque los detalles se conocen mejor que los de la mayoría de los demás núcleos y zonas corticales. Una neurona individual participa en muchos patrones globales como el descrito y tiene escasa significación cuando se la toma aisladamente. En este sentido, se puede decir que el mecanismo básico de reconocimiento visual es la emergencia de un estado global entre *conjuntos neuronales resonantes*. De hecho, Stephen Grosberg ha sido pionero de un análisis detallado de esas redes neuronales reso-

45 Para un examen detallado de esto en el caso de la rivalidad binocular, véase F. Varela y W. Singer, "Neuronal dynamics in the cortico-thalamic pathway as revealed through binocular rivalry", *Experimental Brain Research* 66 (1987): págs. 10-20.

46 W. Singer, "Extraretinal influences in the geniculate", *Physiology Review*, 57, págs. 386-420.

nantes.[47] Estos modelos son interesantes porque concuerdan con la arquitectura general de los caminos visuales que acabamos de delinear, al tiempo que son matemáticamente precisos, con lo cual permiten una simulación e implementación artificial. Por ejemplo, se ha demostrado que estas redes resonantes son capaces de aprender rápidamente a clasificar letras en categorías, sin necesidad de recibir una lista predefinida.

Los símbolos abandonan la escena

Este nuevo enfoque —conexionismo, emergencia, autoorganización, asociación, dinámica de red— es joven y variada. La mayoría de los que se adherirían a ella sostienen opiniones muy diversas acerca de las ciencias cognitivas y su futuro. Teniendo en cuenta esta advertencia, podemos presentar las respuestas que esta perspectiva da a las preguntas que antes planteamos al cognitivismo:

Pregunta 1: ¿Qué es la cognición?
Respuesta: La emergencia de estados globales en una red de componentes simples.

[47] Stephen Grosberg, *Studies in Mind and Brain* (Boston: D. Reidel, 1984). Para una actualización reciente de esta idea, véase G. Carpenter y S. Grosberg, *Computer Graphics and Image Processing*, 37, 1987, págs. 54-115.

Pregunta 2: ¿Cómo funciona?
Respuesta: A través de reglas locales que gobiernan las operaciones individuales y de reglas de cambio que gobiernan la conexión entre los elementos.

Pregunta 3: ¿Cómo saber si un sistema cognitivo funciona adecuadamente?
Respuesta: Cuando vemos que las propiedades emergentes (y la estructura resultante) se corresponden con una aptitud cognitiva específica: una solución feliz para la tarea requerida.

Uno de los aspectos más interesantes de este nuevo enfoque de la ciencias cognitivas es que los símbolos, en sentido convencional, no desempeñan ningún papel. En el enfoque conexionista, la computación simbólica es reemplazada por operaciones numéricas, por ejemplo, las ecuaciones diferenciales que gobiernan un sistema dinámico. Estas operaciones son más afinadas que las realizadas usando símbolos: en un modelo conexionista, el resultado de una sola computación simbólica discreta se obtendría a través de una gran cantidad de operaciones numéricas que gobiernan una red de unidades simples. En tal sistema, los ítems significativos no son símbolos sino complejos patrones de actividad entre las muchas unidades que constituyen la red.

Este enfoque no simbólico implica un radical abandono del supuesto cognitivista básico de que tiene que haber un claro nivel simbólico en

la explicación de la cognición. Como expusimos antes, el cognitivismo utiliza los símbolos para satisfacer la necesidad de un nivel semántico o representacional que sea de naturaleza física. Los símbolos son significantes y físicos a la vez, y el ordenador es un aparato que respeta el sentido de los símbolos pero sólo manipula su forma física. Esta separación entre forma y sentido fue el golpe maestro que dio origen al enfoque cognitivista, y en verdad a la lógica moderna. Pero esta maniobra fundamental también implica una flaqueza cuando se encaran los fenómenos cognitivos en un nivel más profundo: ¿cómo *adquieren* los símbolos su sentido?

En situaciones donde el universo de elementos representables es limitado y definido (por ejemplo, cuando se programa un ordenador o cuando se dirige un experimento con un conjunto de estímulos visuales predefinidos), el origen del sentido es claro. Cada elemento físico o funcional y particular se tiene que *corresponder* con un elemento externo (su referencia) mediante una función que el observador provee fácilmente. Si eliminamos tales restricciones, sólo queda la forma de los símbolos, tan vacía de sentido como un grupo de bits en un ordenador cuyo manual de instrucciones hemos perdido.

Sin embargo, en el enfoque conexionista, el sentido no está localizado en símbolos particulares; está en función del estado global del sistema y está enlazado con el desempeño general en un área determinada, como el reconocimiento o el

aprendizaje. Como este estado global emerge de una red de unidades que son más densas que los símbolos, algunos investigadores se refieren al conexionismo como el "paradigma subsimbólico".[48] Argumentan que los principios formales de la cognición residen en este dominio subsimbólico, un dominio que está por encima de lo biológico, pero también más cerca de lo biológico que el nivel simbólico del cognitivismo. En el nivel subsimbólico, las descripciones cognitivas están elaboradas a partir de componentes que en un nivel superior llamaríamos símbolos discretos. El significado, sin embargo, no reside en estos componentes *per se*, sino en complejos patrones de actividad que emergen de las interacciones de dichos componentes.

Esta diferencia entre lo subsimbólico y lo simbólico nos lleva de vuelta a nuestra pregunta acerca de la relación entre diversos niveles de explicación en el estudio de la cognición. ¿Cómo podrían estar relacionadas la emergencia subsimbólica y la computación simbólica?

La respuesta más obvia es que se trata de dos enfoques, uno ascendente y otro descendente, o que se deberían unir pragmáticamente de un modo mixto, o que simplemente se deberían usar en diferentes niveles o etapas. Un ejemplo típico consistiría en describir las primeras fases

48 Paul Smolensky, "On the Proper Treatment of Connectionism", *Behavior and Brain Sciences*, 11 (1988).

de la visión en términos conexionistas, hasta, por ejemplo, la corteza visual primaria. Luego, en el nivel de la corteza inferotemporal, la descripción se basaría en programas simbólicos. La situación conceptual de tal síntesis, sin embargo, dista de ser clara, y todavía nos faltan ejemplos concretos.

A nuestro juicio, la más interesante relación entre la emergencia subsimbólica y la computación simbólica es una relación de *inclusión*, en la cual vemos los símbolos como una descripción más elevada de propiedades que en última instancia están encastradas en un sistema distribuido subyacente. El caso del llamado "código" genético es paradigmático, y aquí podemos usarlo como ejemplo concreto.

Los biólogos han pensado durante años que las proteínas están codificadas por los nucleótidos del ADN. Sin embargo, es claro que los tripletes de ADN son capaces de seleccionar adecuadamente un aminoácido en una proteína sólo si están encastrados en el metabolismo de la célula, es decir, en medio de miles de regulaciones enzimáticas en una compleja red química. Es sólo a causa de las regularidades emergentes de tal red en su conjunto que podemos hacer abstracción de este trasfondo metabólico y tratar a los tripletes como códigos para aminoácidos. En otras palabras, la descripción simbólica es posible en otro nivel. Es legítimo tratar tales regularidades simbólicas en sí mismas, pero su status e interpretación son muy distintas que cuando las

encaramos como si fueran independientes del sustrato del cual emergen.[49]

El ejemplo de la información genética se puede transponer directamente a las redes cognitivadas con las cuales trabajan los neurocientíficos y los conexionistas. De hecho, algunos investigadores expresaron explícitamente este punto de vista en tiempos recientes.[50] En la teoría de la armonía de Paul Smolensky, por ejemplo, los "átomos" fragmentarios de conocimiento que rodean los circuitos eléctricos están enlazados por algoritmos estadísticos distribuidos, y así brindan un modelo del razonamiento intuitivo en este dominio. La competencia de este sistema se puede describir como la realización de inferencias basadas en rutas simbólicas, pero el desempeño se encuentra en otro nivel y se realiza sin

[49] Para el distingo entre descripción simbólica y emergente y su explicación en los sistemas biológicos, véase Francisco Varela, *Principles of Biological Autonomy* (Nueva York: Elsevier North Holland, 1979, 1979), capítulo 7, y *Autonomie et Connnaissance* (París: Éditions du Seuil, 1989), capítulo 10; y más recientemente Susan Oyama, *The Ontogeny of Information* (Cambridge University Press, 1985).

[50] Véase W. Daniel Hillis, "Intelligence as an Emergent Behavior; or, The Songs of Eden", *Dedaelus*, invierno 1988, págs. 175-189; y Paul Smolensky, "On the Proper Treatment of Connectionism". En una vena muy diferente, véase Jerome Feldman, "Neural Representation of Conceptual Knowledge", University of Rochester TR 189, 1986. Feldman propone una posición intermedia entre sistemas "puntuados" y distribuidos.

referencia a un intérprete simbólico. Esta idea está retratada gráficamente en la Figura 8.

¿En qué difiere este punto de vista mixto de la concepción cognitivista de los niveles de explicación? La diferencia es bastante sutil y consiste ante todo en un cambio de perspectiva. El punto básico, en el cual convienen todos, es que para formular generalizaciones explicativas necesitamos una taxonomía o vocabulario descriptivo adecuado. El cognitivismo, como hemos visto, se basa en la hipótesis de que esta taxonomía consiste en símbolos. Este nivel simbólico *limita* las

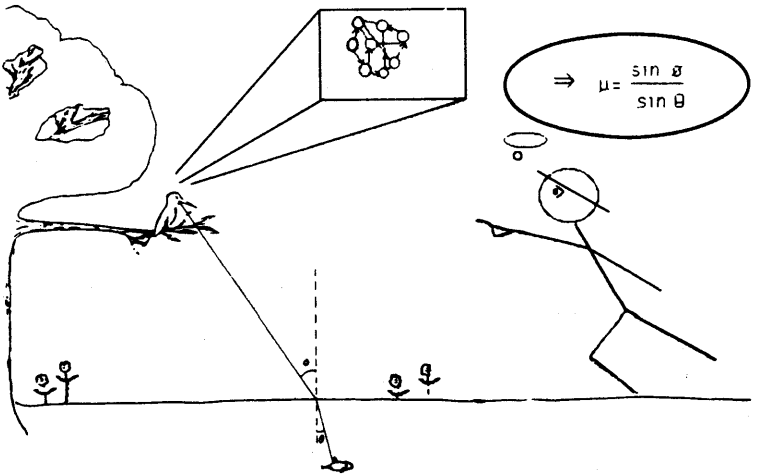


Figura 8. Caricatura de *Punch*: versión modificada. (Véase Figura 3.)

clases de conducta que son posibles para un sistema cognitivo, y así se piensa que tiene una categoría explicativa independiente. En el punto de vista mixto se admite la necesidad de un nivel simbólico, pero queda abierta la posibilidad de que este nivel sea sólo *aproximado*. En otras palabras, los símbolos no se encaran por su valor nominal; se los ve como descripciones aproximadas, en un macronivel, de operaciones cuyos principios rectores residen en un nivel subsimbólico.

Entre los posibles cambios que implica esta síntesis, vale la pena señalar dos en particular. Primero, se puede enfocar con mayor claridad la cuestión del origen de un símbolo y su significado (ejemplo: ¿por qué el triplete ATT selecciona la alanina?). Segundo, todo nivel simbólico pasa a depender de las propiedades y peculiaridades de la red subyacente y queda ligada a su historia. Por lo tanto, una descripción de la cognición basada puramente en procedimientos,* independiente del modo en que la cognición está ligada a su historia, queda en tela de juicio.

El cognitivista sin duda responderá que tal modalidad inclusiva o mixta está bien si uno se interesa sólo en los procesos de nivel "inferior",

* En las CTC, el criterio por procedimientos se opone al declarativo, aludiendo a la descripción del conocimiento por su puesta en obra antes que por sus reglas de producción, lo cual refleja la dicotomía desempeño/competencia a que se alude más arriba. *(N. del T.)*

tales como los hallados en la "codificación" genética. Pero cuando uno encara procesos de nivel "superior", como la capacidad para analizar oraciones o hacer deducciones, se requiere un nivel simbólico independiente. En el caso de las estructuras muy recursivas, tales como el lenguaje humano, se argumentará que el nivel simbólico no es aproximado en absoluto; es la única descripción precisa disponible para formas de representación productivas y sistemáticas.[51]

Hay mucho que decir a favor de este argumento, pero la objeción es que limita injustificablemente el dominio de la cognición a procesos de nivel "superior". Por ejemplo, Jerry Fodor y Zenon Pylyshyn escriben en un artículo reciente: "No sería desatinado describir la ciencia cognitiva clásica [cognitivismo] como un intento de aplicar los métodos de teoría de la prueba a las estructuras del pensamiento (y, análogamente, de cualesquiera otros procesos mentales que impliquen inferencias, sobre todo el aprendizaje y la percepción.) No se trata de que las pruebas lógicas sean por sí mismas tan importantes en el pensamiento humano, sino que el modo de enca-

[51] Esta posición está explicada en detalle por Jerry Fodor y Zenon Pylyshyn en su artículo "Connectionism and Cognitive Architecture: A Critical Review", *Cognition*, 1988. Para una posición filosófica a favor del conexionismo, véase Hubert Dreyfus y Suart Dreyfus, "Making a Mind versus Modeling a Brain: Artificial Intelligence Back at a Branchpoint", *Daedalus*, invierno 1988, págs. 15-43.

rarlas da una pista de cómo encarar los procesos cognitivos en general".[52] A pesar de esta última aclaración, sin embargo, la argumentación que luego desarrollan en el artículo parece requerir que la lógica deductiva sea el paradigma del pensamiento humano y al parecer, por ende, de la cognición en general.

No vemos razones para aceptar esta estrecha concepción de la cognición. Hay muchas clases de sistemas —los sistemas de inmunidad, por ejemplo— cuya conducta se debería ver como cognitiva, pero sus aptitudes no abarcan estos rasgos muy sistemáticos y productivos.[53] Cuando ensanchamos nuestra perspectiva para incluir tales formas de conducta cognitiva, se puede encarar la computación simbólica como una forma de cognición estrecha y muy especializada. Aunque se podría considerar que esta forma especializada tiene un alto grado de autonomía (ignorando el sistema más amplio en el cual está encastrada), el estudio de la cognición incluiría, no obstante, sistemas consistentes en muchas redes de procesos cognitivos, quizá cada cual con un dominio cognitivo distintivo y propio.

El cognitivismo, tal vez en su afán de esta-

52 *Ibid.*

53 Véase Francisco Varela, Antonio Coutinho y Bruno Dupire, "Cognitive Networks: Immune, Neural, and Otherwise", en A. Perelson (comp.), *Theoretical Immunology* (Nueva Jersey: Addison-Wesley, 1988), volumen 2, págs. 359-377.

blecerse como un programa de investigación maduro, ha resistido tal perspectiva. La perspectiva emergentista, sin embargo, tanto en su fase temprana de estudio de los sistemas autoorganizativos como en su actual forma conexionista, está abierta para abarcar una mayor variedad de dominios cognitivos. Una modalidad mixta o inclusiva parece ser, pues, la estrategia natural. La asociación fructífera entre un cognitivismo menos ortodoxo y la visión emergentista, donde las regularidades simbólicas emergen de procesos distribuidos paralelos, es una posibilidad concreta, especialmente en IA, donde predomina el pragmatismo técnico. Esta empresa complementaria sin duda producirá resultados visibles y quizá llegue a ser, durante muchos años, la tendencia dominante en las ciencias cognitivas.

5

Cuarta etapa. La enacción: una alternativa ante la representación

Una insatisfacción más profunda

Es tentador pensar que las CTC están hoy divididas en dos facciones en guerra: un paradigma predominante, favorito del mundo académico y las inversiones tecnológicas en gran escala, y un más reducido grupo de investigadores jóvenes y audaces empresas comerciales que se desplazan en otra dirección.

Pero, por dos buenas razones, esta descripción sería adecuada sólo en un nivel superficial. Primero, porque para la mayoría de los investigadores que participan del nuevo enfoque no resulta imposible buscar una síntesis entre cognitivismo y conexionismo: podría tratarse, respectivamente, de un enfoque descendente y un enfoque ascendente. La situación conceptual de tal síntesis no es clara, como hemos dicho en nuestro capítulo anterior. Una buena posibilidad es una suerte de solución de compromiso, una alianza entre un cognitivismo menos ortodoxo que inclu-

yera procesos paralelos distribuidos de bajo nivel provistos por los enfoques autoorganizativos, especialmente en la IA, dada su orientación técnica y su ánimo oportunista. Esta complementación potencial sin duda producirá resultados visibles, y bien podría convertirse en la tendencia dominante en las CTC por muchos años.

Segundo, y aun más importante, dicha descripción es superficial porque en ambas orientaciones (y por ende en una futura síntesis) *aún* faltan algunas dimensiones esenciales de la cognición. Es preciso insistir en una orientación totalmente distinta de las CTC, nacida de una insatisfacción más profunda que la búsqueda del paralelismo distribuido, y más relacionada con los cimientos mismos de los sistemas representacionales. Cabe esperar que esta tendencia emergente, que hoy goza de cierto espacio, no sufra el mismo destino que las primeras ideas autoorganizativas, que hubieron de ser redescubiertas 30 años después.

El redescubrimiento del sentido común

La insatisfacción central de lo que aquí llamamos el enfoque enactivo es simplemente la total ausencia de sentido común que hay hasta ahora en la definición de cognición. Tanto en el cognitivismo como en el conexionismo de la actualidad, el criterio de cognición continúa siendo una representación atinada de un mundo exter-

no que está dado de antemano. Se habla de elementos informativos a ser captados como rasgos del mundo (como las formas y colores), o bien se encara una definida situación de resolución de problemas que implica un mundo también definido.

Sin embargo, nuestra actividad cognitiva en la vida cotidiana revela que este enfoque de la cognición es demasiado incompleto. Precisamente la mayor capacidad de la cognición viviente consiste en gran medida en *plantear* las cuestiones relevantes que van surgiendo en cada momento de nuestra vida. No son predefinidas sino enactuadas: se las hace emerger desde un trasfondo,* y lo relevante es aquello que nuestro sentido común juzga como tal, siempre dentro de un contexto. Estos dos términos, enactuar y hacer emerger, no son por cierto transparentes en este contexto. La intención de este capítulo es explicarlos mediante un examen conceptual y a través de ejemplos específicos.

Antes de embarcarnos en esta empresa, vale la pena insistir en que se trata de una crítica de la

* El neologismo "enacción" traduce el neologismo inglés *enaction*, derivado de *enact*, "representar", en el sentido de "desempeñar un papel", "actuar". De allí la forma "enactuada": traducir "actuada", "representada" o "puesta en acto" habría llevado a confusión. "Hacer emerger" traduce la forma *bring forth*. En este y otros problemas he seguido el criterio de Pierre Lavoie, el traductor francés, quien aclara que su *faire-émerger* es la traducción del alemán *hervorbringen*, término de origen fenomenológico. *(N. del. T.)*

noción de representación como núcleo de las CTC, ya que sólo se puede representar un mundo que está predefinido. Si el mundo en que vivimos va surgiendo o es modelado en vez de ser predefinido, la noción de representación ya no puede desempeñar un papel protagónico. No se debe subestimar la profundidad de los supuestos a que aludimos aquí. Nuestra tradición occidental ha propiciado (con variantes, desde luego) la comprensión del conocimiento como espejo de la naturaleza.[54] Sólo en el trabajo reciente de algunos pensadores europeos (sobre todo Martin Heidegger, Maurice Merleau-Ponty y Michel Foucault) ha comenzado la crítica explícita de las representaciones. Estos pensadores se interesan en el fenómeno de la *interpretación* entendida como la actividad circular que eslabona la acción y el conocimiento, al conocedor y lo conocido, en un círculo indisociable. Con "hacer emerger" nos referimos a esta total circularidad de la acción / interpretación.[55] Más aun, como esta perspectiva

[54] Para un lúcido comentario sobre esto desde una perspectiva anglosajona, véase R. Rorty, *Philosophy and the Mirror of Nature*(Princeton University Press, 1981).

[55] En este aspecto ha sido muy influyente el trabajo de H. G. Gadamer, *Truth and Method* (Seabury Press, 1975). Para una clara introducción a la hermenéutica véase Palmer, *Hermeneutics* (Northwestern University Press, 1979). La redacción de este capítulo debe mucho a la influencia de F. Flores; véase T. Winnograd y F. Flores, *Understanding Computers and Cognition: A New Foundation for Design* (Nueva Jersey: Ablex, 1986).

analítica enfatiza la acción más que la representación, es adecuado llamar enactivo a este enfoque alternativo de las CTC.[56] En el mundo anglosajón, tradicionalmente más cerca del empirismo lógico, estos temas se han ignorado a menudo.

Sin embargo, en años recientes. algunos investigadores de las CTC han presentado propuestas concretas, llevando estas críticas filosóficas al laboratorio para una reevaluación de la IA. Se trata de una divergencia mucho más radical de las que hubo antes dentro de las CTC, pues trasciende los temas discutidos durante los años de formación. Al mismo tiempo, incorpora muchas de las herramientas desarrolladas dentro del contexto conexionista, como pronto veremos.

El problema de la resolución de problemas

Las CTC siempre han supuesto que el mundo se puede dividir en dominios: regiones de elementos discretos y tareas que el sistema cognitivo enfrenta actuando dentro de un "espacio" dado de problemas: visión, lenguaje, movimiento. Es relativamente fácil definir el dominio del ajedrez: en el "espacio" del juego de ajedrez todo consiste en estados posibles. Hay piezas y posiciones en el tablero. Hay reglas para los movi-

56 El nombre dista de estar establecido. Aquí lo sugiero por razones expositivas, hasta que se proponga uno mejor.

mientos y los turnos. Hay límites claramente definidos. Como se indica en la Figura 9, el "mundo del ajedrez" es casi cristalino. Por lo tanto, como se sabe, hay grandes progresos en el campo del ajedrez por computación. En cambio, ha resultado muy infructuoso llevar este enfoque, por ejemplo, al dominio de los robots móviles. Desde luego, aquí también se pueden escoger ítems discretos (tales como las carrocerías, ruedas y ventanillas de un ensamblaje de automóviles). Pero también es evidente que, mientras el ajedrez tiene límites bien trazados, el mundo del movimiento entre objetos no los tiene. Se requiere el uso continuo del sentido común para configurar nuestro mundo de objetos. Por ejemplo, ¿incluiremos a los peatones en nuestro mundo de conductores de automóviles? Es obvio que la respuesta a estas pregunta sólo se puede extraer de un muy difuso trasfondo de consideraciones irremediablemente contextuales: dónde estamos, qué hora es, por qué calle conducimos y demás. Al contrario del dominio del ajedrez, el dominio del conductor se parece más a una detallada curva fractal que a un cristal claramente definido (Figura 9).[57] El significado de una palabra en un lenguaje natural es quizá un buen ejemplo de todos los ítems que habitan nuestro mundo natural: hay que conocer todo el idioma para percibir

[57] Véase P. Biere, "The professor's challenge", *AI Magazine*, invierno 1985, págs. 60-70.

el significado múltiple de una palabra, que a la vez condiciona el significado de todas las demás. La categorización de cualquier aspecto del mundo natural en que vivimos *no* tiene límites preci-

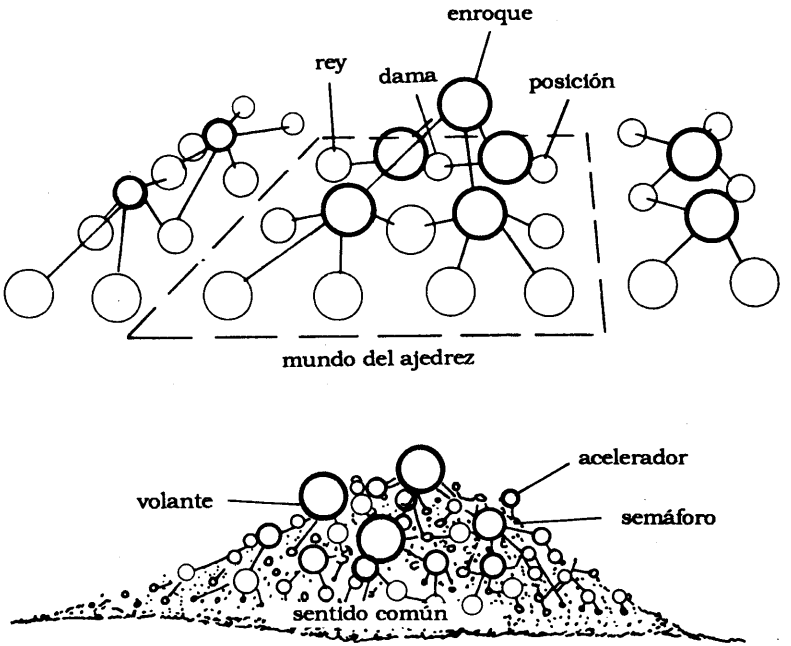


Figura 9. Para el espacio "ajedrez" parece posible diseñar una red de relaciones cuyos nudos representan cada elemento pertinente. En el caso del espacio "conductor automovilístico", una tentativa semejante muestra que, más allá de ciertos elementos aislados, la red evoluciona rápidamente hacia un trasfondo no circunscrito de sentido común. (Fuente: P. Bierre, "The Professor's Challenge", *AI Magazine*, invierno 1985, págs. 60-70.)

sos: no se puede expresar como un dominio a partir del cual elaboramos un mapa.

De hecho, en la década de 1970, al cabo de veinte años de lentos progresos, muchos especialistas en IA advirtieron que la acción cognitiva más simple requiere una cantidad de conocimiento aparentemente infinita, que nosotros damos por sentada (es tan evidente que resulta invisible), pero que se debe servir al ordenador en cucharadas. Como decíamos antes, el gran héroe ha pasado a ser el niño desprejuiciado que aprende a moverse y hablar. A principios de la década de 1960 la investigación se inspiraba en la esperanza cognitivista de hallar un dispositivo *general* para solucionar problemas: una máquina lógica que se pudiera aplicar a *cualquier* problema para resolverlo. Lenta y humildemente, este sueño temprano se redujo a dominios de conocimiento estrictamente *locales* con problemas específicos a resolver, donde el programador pudiera proyectar en la máquina tantos conocimientos de su propia experiencia como fuera posible. Por ejemplo, un sistema experto para reservar billetes en líneas aéreas, pero nada más. Asimismo, la estrategia conexionista exige restringir las clases de atractores posibles a partir de supuestos acerca de las propiedades conocidas del mundo, las cuales se incorporan como reglas adicionales.[58]

[58] Para una descripción concisa y formal, véase T. Poggio, V. Torre y C. Koch, *Nature*, 317:314-319, 1985. El inspirador original fue D. Marrr, *Vision* (Freeman, 1984).

En *ambos* casos, la esquiva ambigüedad del sentido común se deja en la periferia de la investigación, con la esperanza de elucidarla posteriormente.

Estas preocupaciones tienen una desarrollada contrapartida filosófica. Los fenomenólogos se han explayado explicando por qué el conocimiento se relaciona con el hecho de estar en un mundo que resulta inseparable de nuestro cuerpo, nuestro lenguaje y nuestra historia social.[59] Se trata de una interpretación permanente que no se puede aprehender adecuadamente como un conjunto de reglas y supuestos porque es una cuestión de acción e historia; se comprende por imitación, convirtiéndose en parte de una comprensión ya existente. Más aun, no podemos plantarnos fuera del mundo donde nos hallamos para analizar cómo su contenido concuerda con las representaciones: estamos siempre inmersos en él, arrojados en él. Al plantear reglas para explicar la actividad mental y los símbolos como representaciones, nos aislamos precisamente del pivote en torno del cual gira la dimensión viva de la cognición. Ello sólo es posible dentro de un

59 Las principales referencias que tenemos en mente son: M. Heidegger, *Basic Writings* (San Francisco: Harper and Row, 1977); M. Merleau-Ponty, *Phénoménologie de la perception* (Paris: Gallimard, 1976); M. Foucault, *Surveiller et Punir, naissance de la prison* (Paris: Gallimard, 1975); H. Dreyfus, *Why Computers Can't Think* (Nueva York: Macmillan/The Free Press, 1984).

contexto muy limitado donde casi todo permanece constante (los filósofos hablan de una condición *ceteris paribus*). El contexto y el sentido común no son artefactos residuales que se puedan eliminar progresivamente mediante el descubrimiento de reglas más elaboradas. Constituyen la esencia misma de la cognición *creativa*.

Si esta crítica es atinada, no habrá progresos en la comprensión de la cognición en su funcionamiento normal (y no exclusivamente en ámbitos muy limitados) a menos que partamos de otra base que no sea una exterioridad representable.

Las representaciones abandonan la escena

El verdadero desafío que esta orientación plantea a las CTC es que pone en tela de juicio el supuesto más arraigado de nuestra tradición científica: que el mundo tal como lo experimentamos es independiente de quien lo conoce. En cambio, si estamos obligados a concluir que la cognición no se puede entender adecuadamente sin sentido común, el cual no es otra cosa que nuestra historia corporal y social, la inevitable conclusión es que conocedor y conocido, sujeto y objeto, se determinan uno al otro y surgen simultáneamente. En términos filosóficos: el conocimiento es *ontológico*.

En nuestra exposición acerca del cognitivismo distinguimos entre dos sentidos de represen-

tación, y ahora será preciso recordarlos. Por una parte, existe la noción relativamente aceptada de la representación como interpretación: la cognición siempre consiste en interpretar o representar el mundo como si fuera de cierta manera. Por otra parte, existe la idea, mucho más extraña, de que este rasgo de la cognición se debe explicar mediante la hipótesis de que un sistema actúa sobre la base de representaciones internas. Como pareciera que las dos ideas son similares, es preciso afinar un poco esta distinción.

Podemos comenzar señalando un concepto "débil" y admitido de representación. Este concepto es puramente *semántico*: se refiere a todo lo que se pueda interpretar como siendo acerca de algo. Este es el concepto de la representación como "interpretación", pues nada es acerca de otra cosa si no la interpreta de cierta manera. Un mapa, por ejemplo, es acerca de una zona geográfica; representa ciertos rasgos del terreno, y así interpreta que el terreno es de determinada manera. Análogamente, las palabras impresas en una página representan oraciones en un idioma, que a la vez pueden representar —o ser acerca de— otras cosas. Este sentido de representación se puede precisar un poco más. Si, por ejemplo, nos interesan las lenguas en un contexto más formal, podemos decir que los enunciados de un lenguaje representan sus condiciones de satisfacción. Por ejemplo, el enunciado "La nieve es blanca" —tomado literalmente— queda satisfecho si la nieve es blanca; el enunciado "Recoge

tus zapatos" queda satisfecho si la persona interpelada recoge sus zapatos.

Este concepto de representación es "débil" porque no comporta necesariamente ninguna implicación epistemológica u ontológica fuerte. Así es totalmente aceptable decir que un mapa representa el terreno sin preguntarse cómo adquieren los mapas su significado. También es totalmente aceptable pensar que un enunciado representa un conjunto de condiciones sin preguntarse si el lenguaje en conjunto funciona de este modo, o si de veras hay hechos en el mundo separados del lenguaje que luego puedan ser representados por las oraciones del lenguaje. En otras palabras, el concepto débil de representación es pragmático; lo usamos constantemente sin preocuparnos.

La obviedad de tal idea, sin embargo, pronto se convierte en un concepto mucho más fuerte de representación que sí tiene fuertes implicaciones ontológicas y epistemológicas. Este concepto "fuerte" surge cuando generalizamos sobre la base del concepto más débil para elaborar una teoría cabal acerca del funcionamiento de la percepción, el lenguaje o la cognición. Las implicaciones ontológicas y epistemológicas son básicamente dobles: damos por sentado que el mundo está predefinido, es decir, que sus rasgos están definidos antes de toda actividad cognitiva. Luego, para explicar la relación entre esta actividad cognitiva y un mundo predefinido, planteamos la existencia de representaciones mentales

dentro del sistema cognitivo (imágenes, símbolos o patrones subsimbólicos de actividad distribuidos por una red: eso no importa por el momento). Luego tenemos una teoría cabal que establece que: 1) el mundo es predefinido; 2) nuestra cognición aprehende este mundo, aunque sea en forma parcial; y 3) el modo en que conocemos este mundo predefinido consiste en representar sus rasgos y luego actuar sobre la base de estas representaciones.

Para las escuelas representacionistas, toda entidad cognitiva ha caído en paracaídas en un mundo preexistente. Esta entidad sólo sobrevivirá en la medida en que esté provista con un mapa y aprenda a actuar siguiendo ese mapa. En la versión científica de esta historia, el mapa constituye un sistema innato de representaciones (acerca del espacio, el tiempo, las formas, los olores y demás), mientras que aprender a usarlo (patrones motores y emotivos) y actualizarlo (aprender un lenguaje o tarea específica) es tarea de la ontogenia.

Muchos científicos cognitivos replicarán que hemos presentado una mera caricatura. ¿No estamos suponiendo una concepción estática de la representación, una concepción que pasa por alto la riqueza de detalles de la estructura interior de un sistema cognitivo e injustificablemente interpreta que la representación es un mero espejo? ¿Acaso no es bien sabido, por ejemplo, que se considera que la percepción visual es el resultado de un ordenamiento de los patrones físicos de energía que

estimulan la retina para crear representaciones de la escena visual, que luego se usan para hacer inferencias y eventualmente para emitir un juicio perceptivo? Se considera que la percepción es un proceso activo de formación de hipótesis, no el simple reflejo de un ámbito dado.

Esta objeción es justa en cierta medida, pero no da en la tecla. No nos proponemos caricaturizar un complejo programa de investigación, sino explicitar algunos supuestos epistemológicos del modo más claro posible. Así, aunque todos convienen —con la posible excepción de algunos conductistas recalcitrantes— en que la representación es un proceso activo, se la sigue concibiendo como un proceso de "recuperación" o "reconstrucción" de rasgos ambientales extrínsecos e independientes. En la visión, por ejemplo, uno habla de "recobrar la forma a partir de la sombra" o "el color a partir del brillo". Se considera que estos rasgos son propiedades extrínsecas del ambiente, las cuales brindan la información necesaria para recobrar propiedades "superiores" de la escena visual, tales como la forma y el color. Persiste la idea básica de un mundo con rasgos predefinidos.[60]

[60] Para esta concepción de la visión, véase David Marr, *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information* (Nueva York: W. H. Freeman, 1982), especialmente la introducción. Para una explicación filosófica de estas ideas, véase Fred I. Dretske, *Knowledge and the Flow of Information* (Cambridge, Massachusetts: Bradford Books/MIT Press, 1981).

Ejemplos de enacción

Tomemos el caso de la visión: ¿qué vino primero, el mundo o la imagen? La respuesta de los investigadores de la visión (tanto cognitivistas como conexionistas) está dada inequívocamente en los nombres de las tareas investigadas: "recobrar la forma a partir de la sombra", o la "profundidad a partir del movimiento", o "el color a partir de diversas iluminaciones". Podemos llamarla la posición de la *gallina*:

> Posición de la *gallina*: El mundo exterior tiene leyes fijas y precede a la imagen que arroja sobre el sistema cognitivo, cuya tarea consiste en aprehenderlo apropiadamente (sea en símbolos o en estados globales).

Esto parece ser muy razonable y resulta muy difícil imaginar que pueda ser de otra manera. Tendemos a pensar que la única otra posibilidad es la posición del *huevo*.

> La posición del *huevo*: El sistema cognitivo crea su propio mundo, y su aparente solidez sólo refleja las leyes internas del organismo.

La orientación enactiva propone un camino intermedio[61] para trascender ambos extremos:

[61] Esto está explicado en mi "Living ways of sense making: A middle way approach to neuroscience", en P. Livingstone (comp.), *Order and Disorder* (Stanford: Anma Libris, 1984).

darnos cuenta de que (como bien saben los granjeros) el huevo y la gallina *se definen mutuamente*, son correlativos. El proceso continuo de la vida ha modelado nuestro mundo en una ida y vuelta entre lo que describimos, desde nuestra perspectiva perceptiva, como limitaciones externas y actividad generada internamente. Los orígenes de este proceso se han perdido para siempre, y en la práctica nuestro mundo es estable (excepto cuando se desmorona). Pero esta aparente estabilidad no tiene por qué obstaculizar una búsqueda de los mecanismos que la hicieron emerger. Lo que marca la diferencia entre el enfoque enactivo y cualquier forma de constructivismo[62] o neokantismo biológico[63] es este énfasis en la codeterminación (del huevo y la gallina). Es importante tenerlo en cuenta, pues la filosofía más o menos realista que impregna las ciencias cognitivas suele dar por sentado que quien cuestione las representaciones debe asumir *ipso facto* la posición antitética, acechada por el espectro del solipsismo. Pero no nos llamemos a engaño creyendo que el conocimiento opera mediante la representación de una aparente exterioridad.

[62] Véase por ejemplo P. Watzlawick (comp.), *The Invented Reality: Essays on Constructivism* (Nueva York: Norton, 1985).

[63] Visto con mayor claridad en la escuela vienesa de Konrad Lorenz, tal como está expresada, por ejemplo, en un trabajo conocido en inglés como *Behind the Mirror* (Harper and Row, 1979), y en francés como *L'Envers du miroir* (París: Flammarion, 1975).

Expresemos las mismas ideas con un ejemplo. Pensemos en los simples sistemas autoorganizativos introducidos en la Figura 6, y quitemos la restricción de que están diseñados para una tarea particular. Dejemos que el sistema simplemente exista dentro de un ámbito que contiene una selección de elementos aleatorios. Luego sigamos su historia y veamos cómo un sistema particular se acopla con este ámbito aleatorio. Como muestra la Figura 10,[64] entre las infinitas secuencias posibles, el sistema cambia su estado global (se desplaza de un atractor a otro) sólo cuando se topa con dos perturbaciones consecutivas. En otras palabras, el sistema ha escogido un subconjunto de acontecimientos (los que consisten en dos perturbaciones sucesivas) como algo significativo, pues sólo ante ese encuentro cambia su configuración de modo regular. En este muy simple ejemplo presenciamos el origen de un mundo para el sistema, que emerge tangiblemente de circunstancias aleatorias a través de una historia de acoplamiento. Desde luego, no hay "representación" de este subconjunto escogido de todas las secuencias aleatorias posibles dentro del sistema, ni era su tarea reco-

[64] Figura tomada de F. Varela, *Autonomie et Connaissance*, *op. cit.*, capítulo 11. Para una exposición detallada, véase F. Varela, "Structural Coupling in a Cellular Automaton", en E. Secarz, F. Celada y M. Mitchinson (comps.), *Semiotics and Cellular Communication* (Berlín: Springer-Verlag, 1988).

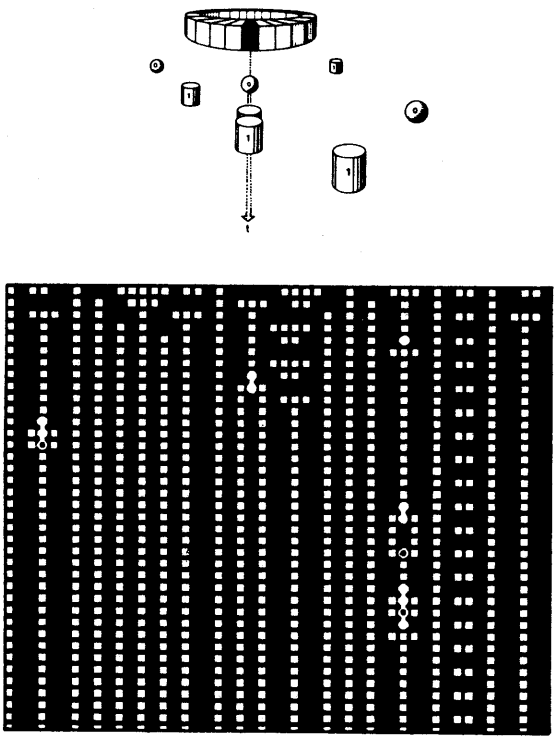


Figura 10. Un anillo —una cadena circular— construido según el principio ilustrado en la Figura 5 está ahora envuelto por un mundo de perturbaciones aleatorias, representadas en la parte superior del diagrama por varios 0 y 1. El anillo entra en contacto con su mundo con cada nueva perturbación, lo cual cambia el estado de una neurona nada. Así, en el decurso del tiempo, una historia del acoplamiento entre anillo y mundo es añadida a la dinámica interna propia del anillo mismo (tal como se lo describe en la Figura 5). En la parte inferior de esta figura mostramos un ejemplo de una tal historia de acoplamiento. Nótese que, en este ejemplo, una o tres perturbaciones sucesivas en una neurona dada no suponen un cambio de una configuración espacial global a otra, mientras que un encuentro con dos perturbaciones sucesivas sí. A través de tal cambio global, el sistema ha escogido una configuración específica de encuentros (es decir, dos perturbaciones sucesivas), cuya relevancia es pues inseparable de la historia de acoplamiento del sistema con este mundo aleatorio.

nocerlas. La existencia del sistema mismo las hizo emerger desde un indefinido trasfondo de posibilidades.

Para tomar un ejemplo más cotidiano, pensemos en el mundo de colores que percibimos todos los días. Sus efectos son tan omnipresentes en nuestra vida que sentimos la tentación de pensar que los colores, tal como los vemos, representan el mundo tal cual *es*. Damos por sentado que el color es un atributo de la longitud de onda de la luz reflejada por los objetos, que nosotros la captamos y la procesamos como información relevante. En rigor, como indican muchas investigaciones, el color percibido de un objeto es en buena medida independiente de la longitud de onda que recibimos.[65] En cambio, hay un complejo proceso (el cual entendemos sólo parcialmente) de comparación cooperativa entre múltiples conjuntos neuronales del cerebro,[66] el cual determina el color de un objeto según el estado cerebral global que corresponde tanto a una imagen de la retina como a cierta expectativa de lo que debería ser dicho objeto.

Así, por ejemplo, si ponemos un papel gris sobre un trasfondo rojo, el papel (físicamente) gris cobra un color verdusco, aunque desde lue-

[65] Véase, por ejemplo, E. Land, *Proc. Natl. Acad. Sci.* (USA), 80:5163-5169, 1983.

[66] P. Gouras y E. Zenner, *Progr. Sensory Physiol.*, 1:139-179, 1981.

go no hay ninguna longitud de onda verde añadida en la región. Este fenómeno se denomina "inducción cromática" y habitualmente se interpreta como una ilusión. Pero esta clase de proceso es el corazón mismo de la visión cromática, pues el color surge cuando un proceso similar se produce en tres clases de células: su actividad relativa especifica qué es el color. Los teóricos modernos hablan pues del color no como representación de la longitud de onda, sino de las propiedades reflexivas de una superficie, pues dicha propiedad es independiente de la iluminación pero más característica del objeto. Aunque esto es sin duda más satisfactorio que considerar el color como longitud de onda, todavía desplaza el problema: ¿cómo define el subsistema, ante todo, aquello que debe reconocer como una superficie? El terco sentido común sale por la puerta pero vuelve por la ventana.

Lo que se puede decir es que nuestro mundo cromático es *viable*: es eficaz, dado que hemos perpetuado nuestro linaje biológico. Las operaciones neuronales cooperativas que subyacen a nuestra percepción del color son resultado de la larga evolución biológica de nuestro grupo de primates. Pero otras especies han creado mundos cromáticos diferentes al realizar operaciones neuronales cooperativas diferentes a partir de sus órganos sensoriales. Por ejemplo, parece que muchos pájaros son tetracromáticos (requieren cuatro colores primarios), mientras que nosotros somos tricrómaticos (nos bastan tres colores pri-

marios).[67] En el dominio del color ni las aves ni nosotros somos más o menos "precisos" en lo que concierne a un dominio que presuntamente es el "mismo", sino que habitamos dos mundos perceptivos de diferentes dimensiones, que por lo tanto no se pueden superponer.

Dicho de otro modo: las muy diferentes historias de acoplamiento estructural de aves y primates han hecho emerger un mundo de datos relevantes que para cada cual es inseparable de su modo de vida. Sólo se requiere que cada senda emprendida sea viable, es decir, que sea una serie ininterrumpida de cambios estructurales. Los mecanismos neuronales que subyacen al color no son la solución de un problema (captar las propiedades cromáticas preexistentes de los objetos), sino el surgimiento simultáneo de la percepción cromática en el hombre o el ave y lo que uno luego puede describir como atributos cromáticos del mundo habitado.

Otra dimensión perceptiva donde se pueden ver estas ideas en juego es el olfato, no a causa de la gama comparativa brindada por la filogenia, sino gracias a modernas técnicas electrofisiológicas. Durante muchos años de trabajo,

[67] Para una reseña de la visión cromática comparativa, véase G. Jacobs, *Comparative Color Vision* (Academic Press, 1983). Para una descripción de la fisiología cromática de las aves, véase F. Varela y otros autores, *Arch. Biol. Med. Exp.*, 16: 291-303, 1983.

Freeman[68] se las ingenió para insertar electrodos en el bulbo olfativo de un conejo, para poder mensurar una pequeña parte de la actividad global mientras el animal se comporta libremente. Descubrió que no existe un claro patrón de actividad global en el bulbo a menos que el animal, entero y despierto, sea expuesto a aromas. Más aun, tales patrones emergentes parecen surgir de un trasfondo de actividad coherente para configurar un atractor coherente. Al igual que el color, el olor no se revela como un mapa pasivo de rasgos externos, sino como la articulación creativa de sentido a partir de lo histórico.[69]

Bajo esta luz, pues, la operación del cerebro se interesa centralmente en la enactuación de mundos a través de la historia de linajes viables: es un órgano que construye mundos en vez de reflejarlos.

Síntesis de la doctrina

La noción básica es que las aptitudes cognitivas están inextricablemente enlazadas con una historia vivida, tal como una senda que no existe pero que se hace al andar. En consecuencia, la

[68] W. Freeman, *Mass Action in the Nervous System* (Academic Press, 1975).

[69] W. Freman y C. Skarda, *Brain Res. Reviews*, 10:145-175, 1985. Significativamente, una sección del artículo se titula: "Una retractación de la 'representación'" (pág. 169).

cognición deja de ser un dispositivo que resuelve problemas mediante representaciones para hacer emerger un mundo donde el único requisito es que la acción sea efectiva: ¿permite la continuidad del sistema involucrado?[70]

Pregunta 1: ¿Qué es la cognición?
Respuesta: Acción efectiva: historia del acoplamiento estructural que enactúa (hace emerger) un mundo.

Pregunta 2: ¿Cómo funciona?
Respuesta: A través de una red de elementos interconectados capaces de cambios estructurales durante una historia ininterrumpida.

Pregunta 3: ¿Cómo saber si un sistema cognitivo funciona adecuadamente?
Respuesta: Cuando se transforma en parte de un mundo de significación preexistente (como lo hacen los vástagos de toda especie), o configura uno nuevo (como ocurre en la historia de la evolución).

Señalemos que en estas respuestas aparecen dos conceptos nuevos, hasta ahora ausentes en las CTC. El primero es que, como las repre-

[70] Para una exposición introductoria completa a este punto de vista, véase H. Maturana y F. Varela, *The Tree of Knowledge: A New look at the biological roots of human understanding* (Boston: New Science Library, 1986).

sentaciones ya no desempeñan un papel central, la *inteligencia* ha dejado de ser la capacidad de resolver un problema para ser la capacidad de *ingresar* en un mundo compartido. El segundo es que el proceso *evolutivo* pasa a reemplazar al diseño orientado hacia tareas específicas. Dicho sin rodeos, así como el conexionismo nació del cognitivismo inspirado por un contacto más estrecho con el cerebro, la orientación enactiva va un paso más allá en la misma dirección para abarcar también la temporalidad del vivir, trátese de una especie (evolución), del individuo (ontogenia) o de la estructura social (cultura).

Trabajar sin representaciones

La búsqueda de modelos no representacionales para el estudio de los fenómenos cognitivos (y ésta es, por cierto, una denominación vaga, al igual que conexionismo) atrae a una cantidad relativamente pequeña de especialistas. Lo interesante es que el número de investigadores que encaran el problema de esta manera no ha cesado de crecer en los últimos años. Además, como explicaré en el capítulo siguiente, muchas de las herramientas del conexionismo tradicional se pueden reformular en este contexto, de modo que las líneas divisorias son tan claras como las que había entre la orientación simbólica y la conexionista.

Es obvio que la estrategia enactiva sólo es po-

sible para la IA si estamos dispuestos a desechar las restricciones operativas de la resolución de problemas específicos predefinidos por el ingeniero. Tal es la intención, por ejemplo, de los llamados "sistemas de clasificación", cuya idea básica consiste en que el sistema enfrenta un ámbito no definido que debe articular en significados.[71] Más generalmente, la simulación de historias prolongadas de acoplamiento y de estrategias evolutivas para descubrir las tendencias básicas donde surge el desempeño inteligente todavía pertenece al futuro. Pero como esta línea de investigación no arroja réditos inmediatos, no suele atraer al competitivo sector de la IA, y por lo tanto se ha trabajado muy poco en esta dirección.

Otra importante zona de influencia del enfoque enactivo está en el área de la informática y del lenguaje. En esta perspectiva, la actividad de la comunicación no consiste en la transferencia de información del emisor al receptor. La comunicación se convierte en la modelación mutua de

71 Véase J. H. Holland, Informe técnico, Universidad de Michigan, 1984. Para otras exposiciones de estas ideas (no necesariamente con el enfoque propuesto aquí), véase J. D. Farmer y N. Packard (comps.), *Evolution, Games and Learning: Models for adaptation in machines and nature*, *Physica D*, 1986. Para el modo en que el sistema de inmunidad puede inspirar investigaciones similares, véase F. Varela, V. Sánchez-Leighton y A. Coutinho (1988), Adaptive strategies gleaned from immune networks, *Theoretical Biology*, B. Goodwin & P. Saunders (comps.) (Edinburgh University Press, 1989).

un mundo común a través de una acción conjunta: el acto social del lenguaje da existencia a nuestro mundo. Hay algunas dimensiones del lenguaje a través de las cuales se hace emerger lo social. Se trata de los actos lingüísticos que realizamos constantemente: declaraciones, promesas, requerimientos y formulaciones. De hecho, dicha red de actos de habla, con sus condiciones de satisfacción, no constituye una mera herramienta para la comunicación sino la trama de nuestra identidad. Da cuenta de mucho de lo que ocurre en una oficina, y por tanto se relaciona mucho más con la comprensión de su dinámica que los clásicos organigramas administrativos.[72] Este enfoque ha originado una nueva rama de herramientas IA conocidas como *coordinadores*.[73] La idea básica es que un coordinador reemplaza el correo electrónico estándar con *software* destinado al análisis de flujo de las conversaciones que se desarrollan en una empresa. Los coordinadores constituyen un ejemplo de reorientación de la IA: en vez de exigir al ordenador lo que quizá sea imposible, es decir, la tarea abierta del análisis lingüístico, se lo usa como un vehículo estructurado para el reconcimiento explícito de nuestra continua enacción lingüística.

[72] Para un comentario sobre este punto de vista, véase L. Smirch y C. Stubbart, *Acad. Manag. Rev.*, 10:724-736, 1985.

[73] T. Winograd y F. Flores, *Understanding Computers and Cognition*, *op. cit.*

Eslabones entre la emergencia y la enacción

Todo eslabón entre la emergencia y la enacción depende de las funciones que atribuyamos a un sistema distribuido. Si enfatizamos un proceso histórico que conduce a regularidades emergentes sin una restricción final fija, recobramos la condición biológica, de final más abierto. Si en cambio enfatizamos una red que adquiere una capacidad muy específica en un dominio muy definido, vuelven las representaciones, y tenemos el más habitual enfoque de los modelos conexionistas. Sin embargo, la primera interpretación también implica una nueva perspectiva sobre la cognición, como explicamos en el capítulo anterior.

En consecuencia, el camino que tomemos dependerá en gran medida de nuestro interés en permanecer más cerca de la realidad biológica, y más lejos de consideraciones pragmáticas. Desde luego, es posible definir un dominio fijo dentro del cual puede funcionar un sistema conexionista, pero ello oculta los más profundos problemas del origen, tan cruciales para el enfoque enactivo.

Pensemos, por ejemplo, en la teoría de la armonía de Smolensky. Su enfoque de la computación subsimbólica como modelo de la intuición parece muy emparentado con la perspectiva enactiva, por lo cual puede ser el mejor ejemplo para utilizarlo como contraste. Sin embargo, aun la teoría de la armonía se evalúa en referencia a

un nivel inviolado de realidad ambiental: los rasgos exógenos concuerdan con ciertos datos del mundo, y una actividad endógena que adquiere, a través de la experiencia, un estado de significación abstracta, una "codificación óptima de la regularidad ambiental". La meta consiste en hallar una actividad endógena que se corresponda con una "caracterización óptima" del medio ambiente.[74] La perspectiva enactiva requeriría que se lleve esta clase de sistema cognitivo a una situación donde lo endógeno y lo exógeno se definan mutuamente a través de una prolongada historia que requeriría sólo un acoplamiento viable, y que ignorara toda forma de adecuación óptima.[75]

Desde luego, esta orientación parece infructuosa desde el punto de vista de la IA, cuya orientación pragmática tiene como objetivo la producción de un sistema que trabaje en algún dominio con poca demora. Mi argumento es que las propiedades cognitivas emergieron de los sistemas vivientes sin tal preocupación por la caracterización óptima. Resultan de historias de compensaciones viables que crean regularidades,

74 P. Smolensky, *op. cit.*, pág. 260.

75 Vale la pena señalar que se pueden aplicar argumentos similares al pensamiento evolutivo de hoy. Para los paralelismos entre representacionismo cognitivo y adaptacionismo evolutivo, véase F. Varela, en P. Livingstone (comp.), *op. cit.* Para una excelente exposición de los mismos problemas adaptados a la evolución y el desarrollo, véase S. Oyama, *The Ontogeny of Information* (Cambridge University Press, 1985).

pero de ninguna manera es obvio que ellas correspondan a un referente único.

Hay pues una tensión entre dos mundos paralelos de la investigación, donde la decisión de estar en contra o a favor de la crítica enactiva está influida tanto por las complejidades de un viraje conceptual como por el mundo tecnológico, cuyas exploraciones están encorsetadas por la camisa de fuerza de la aplicación inmediata. Sospecho que esta tensión se resolverá mediante una creciente brecha entre los componentes científicos y los tecnológicos de las CTC.[76]

[76] Véanse también las observaciones de Roger Schank en *AI Magazine*, otoño 10985, págs. 122-135.

6

Conclusiones

Empezamos en el núcleo de las CTC para desplazarnos hacia lo que podríamos considerar su periferia, es decir, las consideraciones acerca del contexto circundante, y los efectos de la historia biológica y cultural en la cognición y la acción. Desde luego, quienes se aferran a las representaciones como idea clave, ven estas preocupaciones como meras desviaciones temporarias respecto del más preciso reino de la resolución de problemas, que parece más accesible; otros llegan al extremo de afirmar que esos aspectos "vagos" y "filosóficos" no deberían tener lugar en las ciencias cognitivas.

Algunos contrastes que crean estas tensiones se pueden definir de esta manera:

De:	*Hacia:*
tareas específicas	creativo
resolución de problemas	definición de problemas
abstracto, simbólico	ligado a la historia, al cuerpo
universal	contextual
centralizado	distribuido
secuencial, jerárquico	paralelo
mundo predefinido	mundo enactuado
representación	acción productiva
desarrollo por diseño	desarrollo por estrategias evolutivas

Como resumen visual de esta presentación, he bosquejado las tres principales orientaciones comentadas aquí en el mapa polar de la Figura 11. Opino que estas tres sucesivas olas para comprender la cognición y su origen se relacionan entre sí por imbricación sucesiva, como cajas chinas. En la dirección centrípeta, vamos de lo emergente a lo simbólico excluyendo la base de donde emergen los símbolos, y trabajando con los símbolos según su valor nominal. También podemos ir de la enacción al conexionismo estándar suponiendo regularidades dadas en el dominio en el que opera el sistema (es decir, una función de adecuación en un dominio). En la dirección centrífuga excluimos gradualmente lo que parece estable y regular para enfatizar el análisis del origen de tales regularidades, incluidas las dimensiones perceptivas de nuestro mundo humano. Por ello mismo, los conceptos enumerados en la tabla anterior no se deben ver

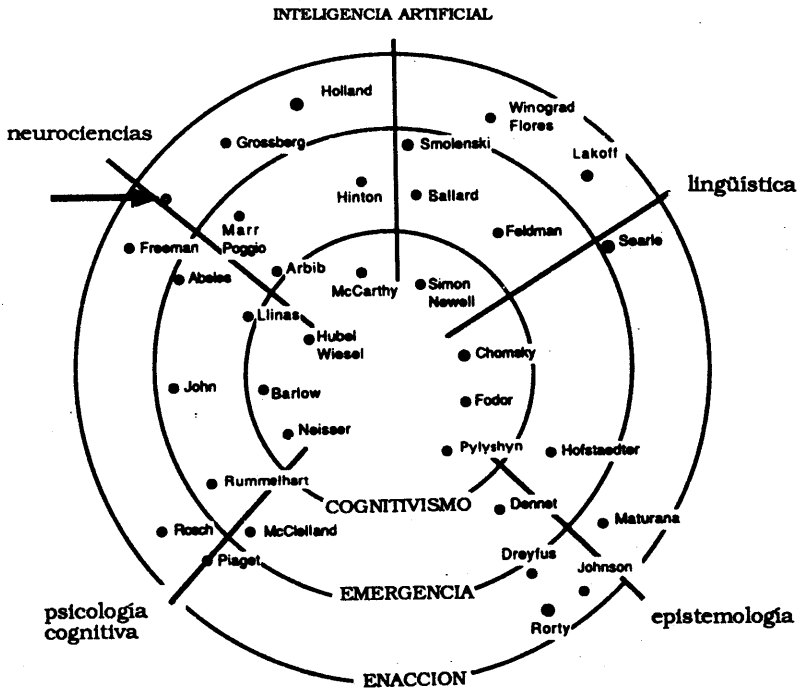


Figura 11. Un mapa polar de las CTC, con el paradigma cognitivista en el centro, los nuevos enfoques en la periferia, y el campo intermedio de las ideas conexionistas entre ambos. El nombre de los investigadores representativos citados en el texto aparece en cada región a lo largo del rayo correspondiente a su disciplina. La flecha indica la posición del autor de este libro.

como opuestos lógicos (o dialécticos). Representan más bien lo particular y lo general, la categoría local y la más abarcadora.

Es obvio que cada uno de estos enfoques, en cuanto niveles de descripción, son útiles en su propio contexto. Sin embargo, si nuestra tarea consiste en comprender el origen de la percepción y la cognición tal como las encontramos en nuestra historia vivida real, creo que el nivel correcto de explicación es el anillo exterior del mapa, el más abarcador. Más aun, para una IA en donde la inteligencia de las máquinas les permita desarrollar un sentido común con los seres humanos, tal como los animales, no veo otro camino que *elevarlas* a través de un proceso de transformaciones evolutivas tal como el sugerido por la perspectiva enactiva. La fertilidad, la dificultad o la imposibilidad de ello es imposible de prever.

En el texto he explicado claramente mis preferencias personales. Ante todo quería demostrar que si el pivote de la cognición es su capacidad para hacer emerger significados, la información no está preestablecida como orden dado, sino que implica regularidades que emergen de las actividades cognitivas mismas. Este reacomodamiento tiene múltiples consecuencias científicas, técnicas, filosóficas y éticas, que ya deberían ser evidentes. En particular, implica una perspectiva para desarrollar las ciencias cognitivas en Europa de una manera singular y vigorosa que sacará el máximo partido de sus singulares tradiciones.

Editorial Gedisa ofrece
los siguientes títulos de la

COLECCION
EL MAMIFERO PARLANTE

Esta colección editorial, dirigida por Eliseo Verón, consta de las obras de los más prestigiosos autores de todo el mundo, y de un amplio espectro de temas, dividida en **Serie Mayor** —dedicada a temas teóricos, técnicos y profesionales— y **Serie Menor** —para un público más amplio—.

Serie Mayor

ELISEO VERÓN	*La semiosis social*
ALAIN BERRENDONNER	*Elementos de pragmática lingüística*
PAUL YONNET	*Juegos, modas y masas*
PIERRE BOURDIEU	*Cosas dichas*
PAUL WATZLAWICK (COMP.)	*La realidad inventada*
GREGORY Y MARY CATHERINE BATESON	*El temor de los ángeles*
FRANCISCO J. VARELA	*Conocer*

Serie Menor

MARC AUGÉ	*El viajero subterráneo*
MARC AUGÉ	*Travesía por los jardines de Luxemburgo*
ELISEO VERÓN	*Construir el acontecimiento*
ISAAC JOSEPH	*El transeúnte y el espacio urbano*

NOTA FINAL

Le recordamos que este libro ha sido prestado gratuitamente para uso exclusivamente educacional bajo condición de ser destruido una vez leído. Si es así, destrúyalo en forma inmediata.

Súmese como voluntario o donante, para promover el crecimiento y la difusión de la Biblioteca

Borre el libro una vez leído

Para otras publicaciones visite
www.lecturasinegoismo.com
Referencia: 795

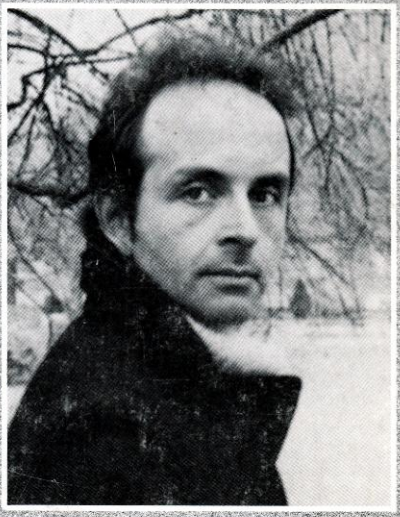
(viene de la cubierta)

siguiente: mientras durante milenios la gentes tuvieron una comprensión espontánea de sí mismas, según la cultura de su época, por primera vez esta visión popular de la mente entra en contacto con la ciencia y es transformada por ella. Muchos deplorarán esta revolución, mientras que otros la celebrarán.

Hoy emerge un nuevo continente del conocimiento, el de las ciencias cognitivas. En la intersección de la informática, la neurobiología y la psicología se construye un enfoque unificado de los fenómenos de la percepción, el (re)conocimiento y la comprensión. El funcionamiento de la mente humana, la conducta animal y el desempeño de los ordenadores son, pues, analizados en una perspectiva común.

El interés conceptual de estas investigaciones, así como la importancia tecnológica de sus aplicaciones, constituyen sin duda el mayor desarrollo científico de este fin de siglo.

En una magistral y concisa síntesis, Francisco Varela traza aquí un panorama de las ciencias cognitivas. A partir de la joven pero compleja historia de estas disciplinas, el autor analiza sus perspectivas actuales y comenta las grandes corrientes ortodoxas que recorrieron este dominio.



El autor

Francisco J. Varela nació en Chile en 1946 y se doctoró en biología en la Universidad de Harvard. Sus trabajos en neurobiología, biología teórica y epistemología, realizados en América latina, los Estados Unidos y Europa, le han granjeado un renombre internacional. Actualmente ocupa la cátedra de epistemología y ciencias cognitivas del Centro de Investigación de Epistemología Aplicada de la Escuela Politécnica de París, y es miembro del Instituto de Neurociencias de la Universidad de París.

gedisa editorial

ISBN 84-7432-383-5

9 788474 323832